# A Behavioral Rebound Effect

Zack Dorner*[1,2,3]

[1]Department of Economics, Monash University, Melbourne, Australia

[2]School of Accounting, Finance and Economics, University of Waikato, Hamilton, New Zealand

[3]Cooperative Research Centre for Water Sensitive Cities, Melbourne, Australia

May 25, 2018

*zack.dorner@waikato.ac.nz.
Supervision provided by Lata Gangadharan, Anke Leroux and Paul Raschky is gratefully acknowledged. Funding was provided by the Commonwealth of Australia through the Cooperative Research Centre for Water Sensitive Cities (CRC grant number 20110044), as well as the Department of Economics in the Monash Business School. This research was made possible by an Australian Government Research Training Program Scholarship. This project has been approved by the Monash University Human Research Ethics Committee; MUHREC#: CF15/3056 - 2015001290. Any errors or omissions are the responsibility of the author alone.

# A Behavioral Rebound Effect


**Abstract**

Behavioral nudges have received significant attention as a potentially cost-effective method to increase energy conservation. Indeed, energy savings from technological improvements can be costly, and are partially offset by the direct rebound effect, whereby a consumer rationally responds to an increase in energy efficiency by consuming more of the energy good. This paper investigates whether technological improvement might also reduce behaviorally motivated energy conservation. A behavioral rebound effect operates through two channels. First, pro-environmental effort is reduced after an increase in energy efficiency. Second, moral licensing reduces pro-environmental effort further when technological change is endogenous. I develop a novel laboratory experiment to exogenously identify these behaviors. I estimate a behavioral rebound effect of 32%. I also find evidence of moral licensing, which is strongest among subjects with a higher degree of pro-environmental attitudes and beliefs. Subjects' baseline level of pro-environmental effort is driven by beliefs about social norms.



Keywords: Energy conservation, energy efficiency, environmental externality, laboratory experiment, moral licensing, pro-environmental behaviors, rebound effect, social norms.

JEL classification: D62, D64, Q40, Q55.

# 1 Introduction

Behavioral nudges provide a potentially powerful and cost effective avenue for decreasing energy use and encouraging pro-environmental behaviors (Allcott and Mullainathan, 2010; Schubert, 2017). For example, environmental campaigners, corporations, governments and economists all understand that individuals have pro-environmental preferences and a proclivity to follow social norms, both of which can lead to pro-environmental choices and behaviors (Cason and Gangadharan, 2002; Croson and Treich, 2014; DEFRA, 2008). Naturally, technological improvement is also vital for addressing environmental issues ranging from water shortages to climate change (Duarte et al., 2014; The Global Commission on the Economy and Climate, 2014). Environmental policy interventions involving behavioral nudges have generally been assumed to be additional to technological improvement, being promoted as a further range of policy levers that can be considered alongside investment in new technology (Allcott and Mullainathan, 2010). However, the impact of technological change on pro-environmental behaviors has yet to be explored, leaving a gap in what might be a more nuanced story.

First, the reduction in environmental damage from technological improvement is partially offset by the rebound effect, whereby a consumer rationally responds to an increase in energy efficiency by consuming more of the energy good. This effect has long been recognised within economics (Jevons, 1865) and highlighting it is not the contribution of this paper. The rebound effect is modeled in the literature as a result of standard income and substitution effects (Chan and Gillingham, 2015). However, there exists a behavioral route by which technological change could lead to a rebound, through its effect on the incentives to undertake pro-environmental behaviors. For example, moving to a more efficient car decreases the relative environmental benefit of walking and cycling, thus reducing the pro-environmental incentives for not driving. Additionally, moral licensing (defined below) could increase the size of such a behavioral rebound whenever a technology improvement has been consciously chosen by the individual. If there is a behavioral rebound effect, then the implication is that behavioral interventions need to be considered within the context of technological change.

The main aim of this paper is to investigate the existence of a behavioral rebound effect and whether improvements in energy efficiency are also subject to moral licensing behavior. I define the behavioral rebound effect as a decrease in pro-environmental effort after an increase in energy efficiency. Pro-environmental effort refers to effort undertaken purely for environmental reasons, such as any walking or cycling done purely on environmental grounds and not for other benefits from these modes of transport, like fitness, enjoyment or saving money. As with the standard rebound effect, the behavioural rebound effect is defined in relation to an exogenous change in energy efficiency. Moral licensing accounts for any additional reduction in pro-environmental effort due to an endogenous change in energy efficiency. Moral licensing is a behavioral phenomenon whereby individuals who undertake a moral action will subsequently behave in an immoral or unethical way (Blanken et al., 2015). Returning to the transport example, after an individual purchases a highly efficient car due to its environmental credentials, they may feel they have a license to no longer walk and cycle for certain trips. Thus, moral licensing has the potential to increase the size of the behavioral rebound effect in the presence of an endogenous increase in energy efficiency.

I develop a novel laboratory experiment to investigate the behavioral rebound effect and moral licensing. The experiment can cleanly isolate pro-environmental behaviors without the many confounds potentially present in the field, such as other motivations to improve energy efficiency or reduce energy usage like saving money. Subjects must decide how to allocate their effort, in a real effort task, between earning money for themselves and avoiding damages to a tree planting charity. I find pro-environmental effort does change with exogenous changes in pro-environmental incentives, and thus there is a behavioral rebound effect. I also find evidence for moral licensing when environmental incentives change endogenously, particularly for individuals with a stronger pro-environmental orientation of their attitudes and beliefs. Finally, the main driver of pro-environmental effort is beliefs about social norms.

There is a significant literature on pro-environmental behaviors, and how they are driven by preferences and social norms (eg. Costa and Kahn, 2013; Croson and Treich, 2014; Sturm and Weimann, 2006). Allcott and Mullainathan (2010) point to the power of non-price, or behavioral, interventions in decreasing energy use, compared with improvements in energy efficiency. This paper adds an important new contribution to the empirical literature by looking at resource conservation from the opposite direction, namely the behavioral implications of technology change. To further contribute to this literature, I also measure drivers of underlying willingness to sacrifice for the environment, including pro-environmental orientation of values and beliefs about social norms. Additionally, the experimental design itself is an innovation; I am not aware of any similar fully incentivized laboratory experiments that measure responses to a consumption externality, which involves real world environmental damages. Finally, this study provides evidence around moral licensing behavior, for which little work has been done within the field of environmental economics (see Tiefenbeck et al., 2013, as one of the few studies).

In the next section I review some background to this study. Section 3 follows with an outline of the method, starting with a definition of the behavioral rebound effect in relation to the canonical model of the rebound effect, and concluding by describing the experiment and the hypotheses. In Section 4 I present the results, followed by discussion and conclusion sections.

## 2 Background

The existing literature on the rebound effect has identified three levels at which the rebound effect operates - the direct rebound effect, the indirect rebound effect and macroeconomic rebound effects. The direct rebound effect relates to the specific good for which there is an energy efficiency improvement. The direct rebound effect can be defined as the efficiency elasticity of an energy service (Sorrell and Dimitropoulos, 2008). Using the car example, this is the percentage change in kilometers driven divided by the percentage change in energy efficiency of the car. The indirect rebound effect relates to other goods. It is the increase in energy usage from an increase in consumption of other goods after an increase in energy efficiency in one good, which can be modeled as the balancing of income and substitution effects within a consumption bundle (Ghosh and Blackhurst, 2014). Finally, macroeconomic rebounds are due to a reduction of market prices for energy in general equilibrium due to lowered demand after increases in the average level of energy efficiency across the economy. This reduction in market price offsets energy savings as consumption of the energy good is encouraged by the reduction in price (Gillingham et al., 2016). While these latter two types of rebound effects are important for the overall picture, this paper is focused at the level of the direct rebound effect.

I measure just the direct behavioral rebound effect as this type of rebound effect is extremely difficult to measure in the field. Focusing on just the behavioral rebound effect removes potential confounds associated with designing an experiment to also measure the direct rebound effect. Furthermore, direct rebound effects have been estimated in the field for a number of energy-consuming goods, particularly transport and heating. While estimates vary, the average estimated size of the direct rebound effect for household energy services, including driving, tends to be in the range of 5 to 40% (De Borger et al., 2016; Gillingham et al., 2016; Sorrell et al., 2009). It is important to note that the macroeconomic rebound effect could be substantial, with recent dynamic modeling showing backfire is a possibility at the macroeconomic level (Chang et al., 2017).

A range of lab and field experiments have shown individuals will undertake actions for the benefit of others and the public good. Theoretically, intrinsic motivation or environmental preferences can explain some pro-environmental behaviors; other motivations include image, identity and expectations about the motivations and behaviors of others (Ariely et al., 2009; Bénabou and Tirole, 2006, 2011; Bowles and Polania-Reyes, 2012; Brekke et al., 2003; Nyborg et al., 2006). These theories underpin empirical literature on pro-environmental behaviors. This literature includes evidence that many individuals will pay a premium on particular consumer products for their "green" credentials (Croson and Treich, 2014). There is also work explaining effort put into recycling, water use reduction and energy conservation using environmental preferences and social norms (Abbott et al., 2013; Allcott, 2011; Ayres et al., 2013; Costa and Kahn, 2013; Ferraro and Price, 2013; Halvorsen, 2008). Important for this paper is that while heterogeneous, many individuals do exhibit a willingness to make some personal sacrifice for the environment (Sturm and Weimann, 2006). Additionally, the fact that environmental behaviors are heterogeneous means questions of heterogeneity in pro-environmental attitudes and behaviors can be explored even with the standard student subject pool, which is otherwise largely homogeneous.

Moral licensing has the potential to increase the behavioral rebound effect associated with technological change when that change is endogenous. Since the first study identifying moral licensing (Monin and Miller, 2001), the effect has been found in a number of studies, within and between a range of domains. Blanken et al. (2015) undertake a meta-analysis of 91 studies and find a small to medium effect of moral licensing, in comparison with other effect sizes of behavioral patterns within the field of social psychology. Domains studied include job hiring, racist attitudes, charitable donations and consumer behavior. Within environmental economics, Tiefenbeck et al. (2013) find a water conservation campaign in an apartment complex that resulted in a 6% reduction of water use saw electricity use increase by 5.6% for the treatment group, compared with the control group. Moral licensing could increase the rebound effect if an individual purchases a particularly durable good such as a car, and uses this purchase to psychologically justify driving more.

Laboratory experiments have been successfully utilized as a method for gaining greater insight into real world economic behaviors in a range of contexts, including environmental economics (Friesen and Gangadharan, 2013; Sturm and Weimann, 2006). A strength of the method is the high level of control it accords the researcher in measuring very specific treatment effects, with a high degree of confidence in claims of exogeneity and a minimization of potential confounds. This trait makes laboratory experiments particularly suited to investigating behavioral responses to real world phenomena or policies that are difficult to isolate in the field. Limitations of the laboratory environment include the behavioral implications of a high level of salience to subjects of the effect of their actions – in this case environmental damages – and an awareness of being observed (Schubert, 2017; Levitt and List, 2007). Understanding the implications of these limitations has helped guide the experimental design and interpretation of results presented here.

In the case of the rebound effect, behavioral responses to technological change are particularly tricky to identify in the field. Investment by households in durable goods is an endogenous decision, including the choice of level of energy efficiency of a vehicle or appliance (De Borger et al., 2016). Secondary field data has been important in measuring the rebound effect and is indeed the primary means by which the rebound effect is measured. However, for the reasons just mentioned, this is not a straightforward task, meaning there is considerable variance of estimates of the rebound effect in the literature and some methodological debate (Gillingham et al., 2016; Hunt and Ryan, 2014; Sorrell et al., 2009).

Beyond the endeavor of measuring the rebound effect is testing the theory underpinning the hypothesized drivers of the rebound effect. In this case, endogenous investments prove even more problematic to investigating the importance of specific drivers, such as underlying environmental and social preferences and other behavioral phenomena. This is because investment in energy efficiency is likely to be highly correlated with environmental preferences and beliefs about social norms. Research in the lab is a low cost means by which to investigate particular treatment effects, such as response of pro-environmental effort to change in energy efficiency, while ensuring highly credible exogeneity. A carefully considered field experiment into the rebound effect may be highly valuable in this regard too, but a laboratory experiment will increase the evidence base and potentially inform the design for more high cost field work. Thus, a laboratory experiment is highly suited to the research aims of this paper.

# 3 Method

## 3.1 Defining the behavioral rebound effect

I divide this part of the method section into two subsections. First, I discuss the basic definition of the direct rebound effect, given by Sorrell and Dimitropoulos (2008). In the second part I extend the model to include pro-environmental effort to define a behavioral rebound effect.

### 3.1.1 The basic model of the direct rebound effect

Let us define an energy service, $ES$, as $ES = es[S, A]$. $S$ is useful work (in the physics use of the term, such as kilometers traveled) and $A$ is other attributes of the service (for example comfort). In the basic model, useful work is produced from energy through the following relation:

$$S = \epsilon E. \tag{1}$$

The term $\epsilon$ is energy efficiency; effectively it is an output-input ratio, which is a function of capital. $E$ is energy, provided by inputs such as petrol or electricity.[1]

An individual decides on the amount of $S$ to consume, given their preferences, budget constraint and the total cost of consuming $S$. Let $S^*$ be the optimal level of $S$ chosen by the individual. To illustrate how $S^*$ is chosen, let $P_S$ be the price of the energy component of $S$, which is one component of the total cost of consuming $S$. Other components of total cost include maintenance of capital and time costs. These costs are held constant. Thus, the energy cost of $S$ is given by:

$$C_S = P_S S \tag{2a}$$

$$= P_S \epsilon E \tag{2b}$$

$$= P_E E, \tag{2c}$$

where $C_S$ is the energy cost of $S$, and thus $P_E$ is the price per unit of energy. Therefore, as shown above, $P_S = P_E/\epsilon$. This relationship between the change in the energy cost of $S$ and a change in energy efficiency, $\epsilon$, is what the rebound effect hinges on. While a number of variables, including $P_E$, will affect the optimal choice of $S$, $S^*$, all variables are held constant except $\epsilon$ in this analysis. Therefore, I focus on the effect of $\epsilon$ on $S^*$ through the function $S^*(\epsilon)$.

[1] More generally, $E$ could be any resource for which its use is associated with an environmental externality, such as water. However, I keep with the rebound effect literature by calling this resource energy.

With no change in $S^*$ after an increase in $\epsilon$, there is no rebound effect; energy use decreases in proportion to any increase in energy efficiency. However, $S^*$ may increase after an increase in $\epsilon$, holding the price of energy, $P_E$, constant. An increase in $\epsilon$ reduces $P_S$, and thus can increase $S^*$ through positive income and substitution effects (Chan and Gillingham, 2015). In this case, there is a positive rebound effect.

Rearranging equation (1) and taking the derivative of $E$ with respect to $\epsilon$, we get the change in energy use in response to a change in energy efficiency:

$$\frac{\partial E}{\partial \epsilon} = -\frac{S^*(\epsilon)}{\epsilon^2} + \frac{1}{\epsilon}\frac{\partial S^*(\epsilon)}{\partial \epsilon}. \tag{3}$$

Assuming an increase in $\epsilon$, the first RHS term of this equation is the direct change in energy use due to a change in energy efficiency and no change in $S^*$. This term can thus be interpreted as the change in energy use due to simple engineering calculations. The second term on the RHS of the equation is the increase in energy use due to an increase in $S^*$ after an improvement in energy efficiency. Thus, this second term is the increase in energy use due to the direct rebound effect. The size of this term is determined by the size of the income and substitution effects, which both act positively on $E$ (assuming $S$ is a normal good). If there is no rebound effect, there are no income and substitution effects, $S^*$ is no longer a function of $\epsilon$ and this last term in equation (3) falls away.

The direct rebound effect is specifically defined as the proportional increase in useful work from the energy service consumed relative to the proportional increase in energy efficiency. This is equivalent to the efficiency elasticity of demand for useful work:

$$\eta_\epsilon(S) = \frac{\partial S^*(\epsilon)}{\partial \epsilon}\frac{\epsilon}{S^*(\epsilon)}. \tag{4}$$

In the absence of a direct rebound effect, all improvements in energy efficiency lead to a 1 for 1 reduction in energy use. In this case, $\eta_\epsilon(S) = 0$. With a positive direct rebound effect, $\eta_\epsilon(S) > 0$. Backfire occurs when the direct rebound effect is so great that energy usage actually increases after an improvement in energy efficiency, in which case $\eta_\epsilon(S) > 1$.

#### 3.1.2 The behavioral rebound effect

I now extend the basic definitions to include pro-environmental effort, in order to define the behavioral rebound effect. Pro-environmental effort is undertaken to conserve energy purely for environmental reasons, for example riding a bicycle to avoid consuming petrol by driving. Pro-environmental effort is positive when individuals are sufficiently motivated by their pro-environmental preferences or preferences to conform with social norms, given the costs (monetary or otherwise) of undertaking such effort. An important relation underpinning this extended model is the negative association between the effectiveness of pro-environmental effort and energy efficiency, $\epsilon$. When the energy efficiency of a car improves, the amount of energy saved per kilometer by riding a bicycle (instead of driving) falls. Many other pro-environmental behaviors in this example also follow this logic - keeping tires inflated or having a light foot on the accelerator also save less petrol per kilometer with an efficient car compared with an inefficient car.

To consider the extended model more formally, let $M$ be pro-environmental effort. I define $M$ such that it only incorporates effort expended for environmental reasons - either due to pro-environmental preferences or social norms. The term $M$ does not include ostensibly pro-environmental effort, such as riding a bike, where that effort is done to advance other objectives, such as to save money, for enjoyment or to get fit.

Let the energy conserved by pro-environmental effort, $E_M$, be given by:

$$E_M = \phi M. \tag{5}$$

The term $\phi$ is the effectiveness of pro-environmental effort in reducing energy usage, which is an output-input ratio of energy savings from pro-environmental effort. In this extended model, the energy used by consuming $S$, previously defined from equation (1), is given by:

$$E = \frac{S}{\epsilon} - E_M \tag{6a}$$
$$= \frac{S}{\epsilon} - \phi M. \tag{6b}$$

Hence, pro-environmental effort is a substitute for energy, $E$, which is defined as an environmentally damaging energy source, like gasoline. Useful work consumed, $S^*$, is assumed to be a function only of $\epsilon$, and is not affected by pro-environmental preferences or social norms regarding pro-environmental effort. Therefore, $S^*$ in this model can be interpreted as useful work consumed in absence of pro-environmental preferences and preferences to conform to social norms. Optimal level of pro-environmental effort, $M^*$, is a function of $\phi$ as the level of pro-environmental effort depends on the effectiveness of pro-environmental effort, given pro-environmental preferences and social norms, and the private costs incurred from undertaking pro-environmental effort. Preferences and effort costs are held constant.

As noted at the start of this section, $\phi$ is a function of $\epsilon$ such that:

$$\frac{\partial \phi(\epsilon)}{\partial \epsilon} < 0. \tag{7}$$

Thus, an improvement in energy efficiency reduces the environmental benefits per unit of pro-environmental effort.

I now derive equation (3) for the extended model:

$$\frac{\partial E}{\partial \epsilon} = -\frac{S^*(\epsilon)}{\epsilon^2} + \frac{1}{\epsilon}\frac{\partial S^*(\epsilon)}{\partial \epsilon} - \frac{\partial \phi(\epsilon)}{\partial \epsilon} M^*(\phi(\epsilon)) - \phi(\epsilon)\frac{\partial M^*(\phi(\epsilon))}{\partial \phi(\epsilon)}\frac{\partial \phi(\epsilon)}{\partial \epsilon}. \tag{8}$$

Note that $S^*$ and $M^*$ are jointly chosen to maximize the individual's utility, hence I consider their simultaneous effect on change in $E$. The first two terms on the RHS of the equation are unchanged from the base model, however their interpretation changes slightly. The first term on the right-hand side is now just one part of the engineering calculation. The engineering calculation must also include the third term on the RHS of the equation. This term is the change in energy conserved given a change in energy efficiency, but no change in pro-environmental effort.

The second term on the RHS of equation (8) is the change in energy use due to an increase in consumption of $S$; termed the direct rebound effect, as before. This term only incorporates the change in consumption of useful work from the energy service due to private income and substitution effects and does not include pro-environmental preferences or preferences to conform to social norms. The last term on the RHS gives the change in $M^*$ caused by an increase in $\epsilon$, which is a result of the behavioral rebound effect. If $\frac{\partial M^*(\phi(\epsilon))}{\partial \phi(\epsilon)} > 0$, then this final term in equation (8) is positive, hence there is a positive behavioral rebound effect. That is, the change in pro-environmental behaviors leads to less energy savings from an improvement in energy efficiency than predicted solely by the engineering calculations. Therefore, this extended model separates out the direct rebound effect and the behavioral rebound effect. The combination of these two rebound effects determine the overall rebound effect as it pertains to energy use, $E$.

It is important to emphasize that this model hinges on the definition of $M$ as pure pro-environmental effort. In my experiment I can measure pro-environmental effort directly, hence it is useful to separate the direct rebound effect from the behavioral rebound effect. However, in the field it would be difficult to measure $M$ specifically. For example, in the base model, some pro-environmental effort might be captured by a higher $\epsilon$. For example, ensuring tires are fully inflated or using a light foot on the accelerator pedal would improve the observed fuel efficiency of a car. Other pro-environmental effort would be captured through a lower observed $S^*$, such as reducing distance driven. Therefore, this extended model is intended to complement the existing literature on the rebound effect by providing a formulation that allows for behavioral rebounds to be explicitly understood and measured.

Thus, the extended model defines a behavioral rebound effect, equivalent to the negative of the energy efficiency elasticity of pro-environmental effort:

$$-\eta_\epsilon(M) = -\frac{\partial M^*(\phi(\epsilon))}{\partial \epsilon}\frac{\epsilon}{M^*(\phi(\epsilon))} \tag{9a}$$

$$= -\frac{\partial M^*(\phi(\epsilon))}{\partial \phi(\epsilon)}\frac{\partial \phi(\epsilon)}{\partial \epsilon}\frac{\epsilon}{M^*(\phi(\epsilon))}. \tag{9b}$$

Hence, there is no behavioral rebound effect when $-\eta_\epsilon(M) = 0$, a positive behavioral rebound effect when $-\eta_\epsilon(M) > 0$, and backfire when $-\eta_\epsilon(M) > 1$.

Moral licensing has the effect of increasing the size of the behavioral rebound effect. After an individual makes a moral choice, moral licensing is revealed as a subsequent immoral action or a decrease in the level of moral effort the individual otherwise would have made. Thus, if there is a larger behavioral rebound effect after an endogenous increase in $\epsilon$ compared with the same change in $\epsilon$ imposed exogenously, then moral licensing has occurred. This comparison must be undertaken with equal costs for the change in $\epsilon$ to ensure it is not the cost of the choice that is driving the reduction in pro-environmental effort.

The main aim of this experiment is to estimate the behavioral rebound effect, $-\eta_\epsilon(M)$. Through this exercise I can test whether pro-environmental effort, $M$, is an important component of the overall rebound effect in energy use, $E$. The experimental design allows me to estimate the behavioral rebound effect without confounding it with the direct rebound effect. Thus, my experimental design is aimed at measuring just the behavioral rebound effect; it is beyond the scope of this paper to measure the full rebound effect in $E$, in a laboratory setting. By measuring the behavioral rebound effect I can compare its magnitude to the direct rebound effect of energy use as measured in prior research in the field. Given the behavioral rebound effect can also be decomposed into income and substitution effects as they relate to trading off private consumption and reducing environmental damages, I additionally measure just the income effect component of the behavioral rebound effect. Another important component of the behavioral rebound effect is $\eta_\phi(M)$, which is the effectiveness elasticity of pro-environmental effort. This elasticity has a direct impact on the size of the behavioral rebound effect, as follows from equation

(9). Hence I measure $\eta_\phi(M)$ directly, without an associated change in $\epsilon$. Finally, I test whether there are moral licensing effects, which are shown if $-\eta_\epsilon(M)$ with an endogenous increase in $\epsilon$ is greater than $-\eta_\epsilon(M)$ with an exogenous increase in $\epsilon$.

### 3.2 Experimental design

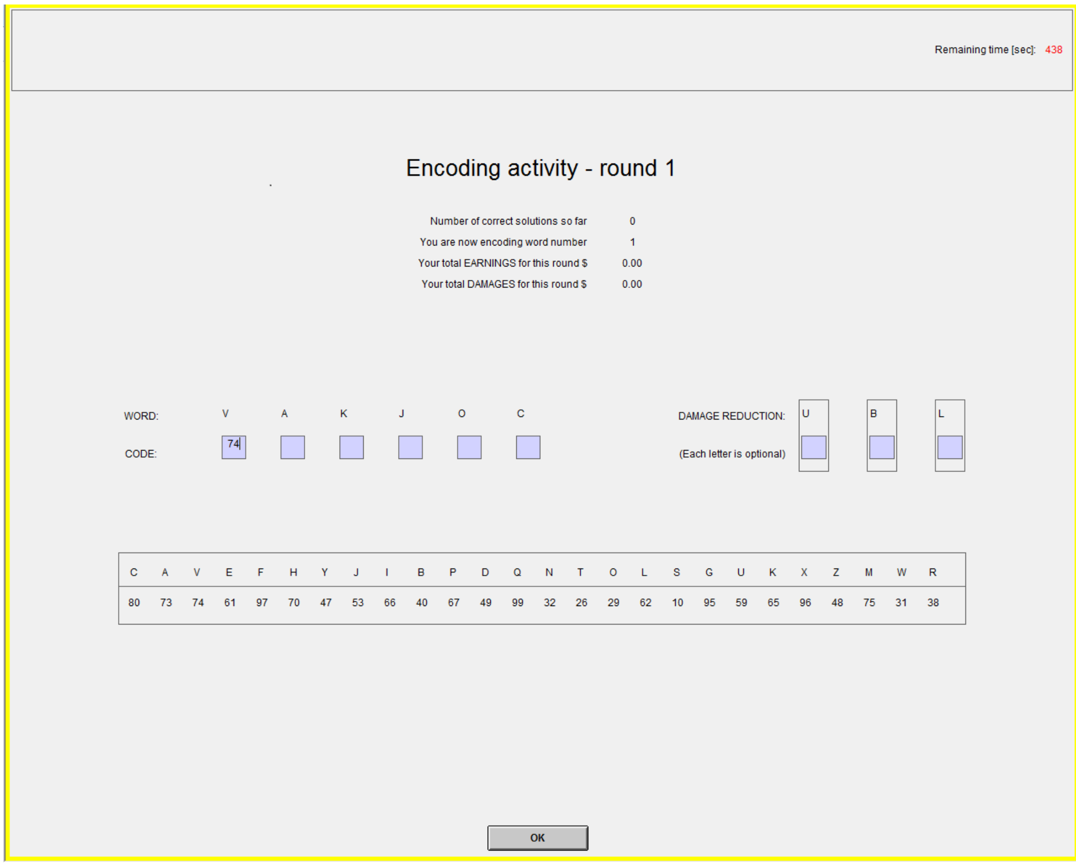
The basic design of the experiment allows the estimation of how subjects trade off between their consumption (monetary earnings) and environmental damage (reduction in a donation to a tree planting charity). By varying damages between rounds (within subjects), I can estimate the size of the behavioral rebound for each subject. By varying the treatments shown to subjects, the experimental design also allows testing between subject hypotheses, such as moral licensing. To link the experiment with the model, monetary earnings before any sacrifice for the environment can be thought of as $S$, environmental damages can be thought of as $E$, and $\epsilon$ determines the level of damages associated with $S$. Damages can be reduced through pro-environmental effort, $M$ at a relative cost of $\phi$.

The experimental activity is based on a word decoding effort task, similar to Erkal et al. (2011) and Benndorf et al. (2018). At the start of each eight minute round, subjects are presented with a screen as shown in Figure 1. Subjects must correctly enter the two digit codes for each of the random letters for the six letter "word" they are given. The codes are provided in a scrambled alphabet across the bottom of the screen. The word is displayed in the center left of the screen. Once a subject has correctly completed the word, she can click the OK button and earn the payment for that word - which is 60c for most treatments.[2] For the subject to maximize her earnings for the round, she must try to complete as many six letter words as possible within the eight minute time limit.

Each completed word reduces a charity payment for that round. The charity is a local tree planting charity, and subjects know that every $2 donated to the charity leads to one seedling being planted. In the high damage treatment, each word completed reduces the charity payment by 54c. However, in the center right of the screen, subjects can lower the damages to the charity for that word by filling in additional letters. It is made clear to the

[2]All monetary values are in Australian dollars.

Figure 1: Experimental screen of the main activity.




subjects that these additional letters are optional, both on screen and in the instructions. One additional letter will lower the damages for that word by one third, two by two thirds and all three additional letters will lower the damages to the charity to nothing. As filling in the additional letters takes extra time, subjects must trade off how much damage they are willing to do to the charity payment (the environment) with their private earnings in each round. Cumulative earnings and damages for the current round are displayed in the top center of the screen. The full instructions are provided in the Appendix.

A real effort task was chosen for the experimental activity given consumption of energy goods and pro-environmental effort in the field both require real effort (the former through earning income and/or through labour input such as driving). The experimental design means there is a clear opportunity cost of private consumption when completing the optional extra letters. This design helps to ensure the measurement variable of interest, pro-environmental effort, is responsive enough to changes in incentives given the sample size, which is not always the case when using real effort tasks (Araujo et al., 2016; Erkal et al., 2017).

The word decoding task is a modified version of Erkal et al. (2011), who did not include optional extra letters. Additionally, the order of the alphabet is scrambled for each word, as suggested by Benndorf et al. (2018). This scrambling is done to minimize any learning effects between rounds and was successful in this case as no learning effect was observed (see Section 4.1). Finally, none of the eight or nine letters given to subjects to decode for each word were repeated within that word. In the piloting stage of the experiment it was observed that subjects were more likely to complete the optional extra letters if they were repeats of letters already given for that particular word.

Each session consisted of 24 subjects. The donation to the charity was made at a session level. The initial donation for any given round was set at \$336; this fact was communicated to subjects in the instructions. This amount meant that there was \$14 donated per subject, which was high enough to ensure the session level charity donation was not depleted to \$0 for a given round, therefore ensuring marginal damages were never 0. Each subject completed a practice round of eight minutes, plus three rounds of eight minutes each. One of the three rounds was paid out.

### 3.2.1 Treatments

There were five treatments overall. Each subject was given one treatment per round, thus each subject saw three treatments. The treatments are shown in Table 1, with experimental parameters and their equivalent term from the theoretical model. Payment per word is equivalent to one unit of $S$ in the theoretical model, when no optional additional letters are completed. As energy consumption is associated with damages, damages per word is equivalent to $E$ consumed per unit of $S$, expressed in terms of environmental damages. The optional extra letters per word provide a maximum level of pro-environmental effort, $M$, that is possible per unit of $S$. Energy efficiency is calculated according to $\epsilon = S/E$, when no optional additional letters are completed, hence it can be calculated by dividing payoff per word by maximum damages per word. Finally, effectiveness of pro-environmental effort, $\phi$, is the ratio of reduction in damages to sacrifice of consumption

Table 1: Treatment parameters.

| *Experimental parameter* | Payoff/word | Damage/word | Optional letters/word | Energy efficiency | Effectiv. of pro-env. effort |
|---|---|---|---|---|---|
| *Theoretical interpretation* | Unit of $S$ | $E$ consumed per unit of $S$ | Max $M$ per unit of $S$ | $\epsilon$ | $\phi$ |
| **Treatment** | | | | | |
| High damage | \$0.60 | \$0.54 | 3 | 1.1 | 2.7 |
| Low damage | \$0.60 | \$0.36 | 3 | 1.7 | 1.8 |
| Choice | \$0.60 | \$0.54 or \$0.36 | 3 | 1.1 or 1.7 | 2.7 or 1.8 |
| Low effort | \$0.60 | \$0.54 | 2 | 1.1 | 3.6 |
| High income | \$0.80 | \$0.72 | 3 | 1.1 | 2.7 |

given $\phi = E_M/M$, as per equation (5).[3] Private earnings for a round, $Y$, for subject $i$ is determined by $Y_i = S_i - M_i$, where $S_i$ is total round earnings absent pro-environmental effort and $M_i$ is earnings sacrificed for the environment.

Running through the treatment values shown in Table 1, high damage, low damage, choice and low effort treatments all pay \$0.60 per word, but vary by damages and number of optional extra letters per word. High damage treatment has damages of \$0.54 per word, whereas low damage has damages of \$0.36 per word. Choice tests for moral licensing - at the start of the round, subjects are given the costless choice of causing either \$0.54 or \$0.36 of damages per word.[4] Low effort tests what happens when $\phi$ is increased without an increase in $\epsilon$. This increase in $\phi$ is achieved by lowering the number of optional extra letters from three to two, where one extra letter completed lowers the damages by half, and two extra letters lowers the damages to 0. Thus, damages per word are the same as high damages, so $\epsilon$ is unchanged, whereas $\phi$ increases. Finally, the high income treatment provides a test of pure income effects - payoff per word and damages per word are both increased by one third relative to high damage.

By design $\phi > 1$ for each treatment to ensure total welfare within the experimental

[3] I provide an example of how $\phi$ is calculated using the high damage treatment (see Table 1). One unit of $M$ is one extra letter, thus in monetary terms it is equivalent to sacrificing 1/6 of the earnings per word, or \$0.10 in the high damage treatment. The damage reduction from one unit of $M$, or $E_M$, is the sum of two values. The first part of $E_M$ is 1/3 of the damages per word, as explained above, or \$0.18 in the case of the high damage treatment. However, it also reduces the amount of words the subject can complete within the eight minute time limit by 1/6. This gives an additional damage reduction of 1/6 of the damages caused by a word with no pro-environmental effort. Hence, for the high damage case, $E_M = \$0.18 + \$0.09 = \$0.27$. Thus, $\phi = 2.7$ for the high damage treatment.

[4] The choice is costless in order to ensure there are no income effects confounding the difference between pro-environmental effort in the low damage treatment and those who chose low damages, as noted in Section 3.1.2.

session (subject payoffs plus donation to the charity) is highest when subjects always complete all optional extra letters. Thus, $\phi$ is akin to the multiplier used in standard experimental games, such as public good and trust games, where donations to a public good or to other players are increased in value by the experimenter (Berg et al., 1995; Sturm and Weimann, 2006). Also note that $\phi = 3/\epsilon$ except for the low effort treatment where $\phi = 4/\epsilon$.[5] Hence, the assumption given in equation (7) holds.

The trade-offs faced by subjects in each treatment can be represented as a budget constraint, as shown in Figure 2. The example shown represents a subject who is able to complete 126 letters within the eight minute time period, and graphs the various allocations of letters between reducing damages to the environment ($E_{Mi}$ on the $x$ axis) and private income ($Y_i$ on the $y$ axis).[6]

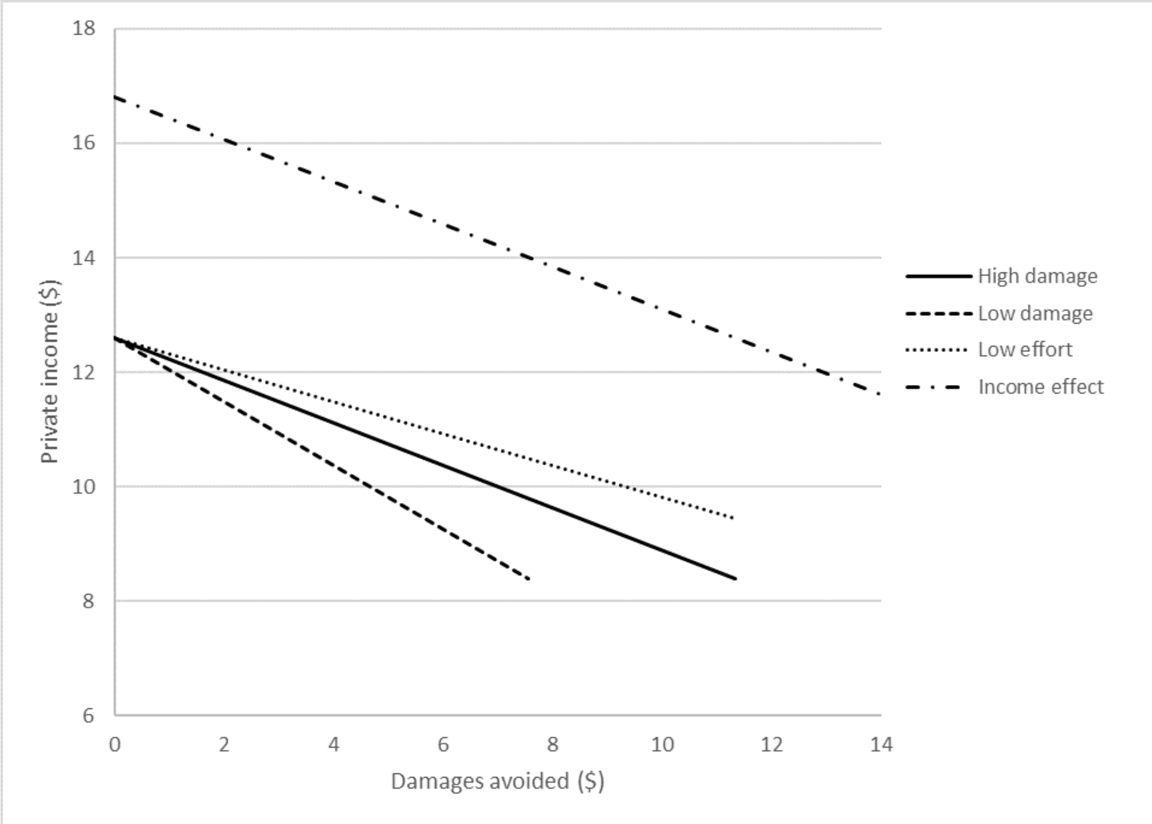
There are five treatment groups, grouped by the treatments and the order of treatments the groups received. These groups are shown in Table 2, along with the number of subjects in each group. These treatment groups allow for the testing of between subject hypotheses using a differences-in-differences approach. Specifically, comparing the difference in pro-environmental effort between treatment groups A and B with C for rounds 1 and 2 allows for the testing of order effects to ensure they do not play a role in driving the overall results. Comparing the difference in pro-environmental effort between rounds 1 and 2 between treatment groups A and B and treatment groups D and E allows for the testing of moral licensing.[7]

[5]These relationships between $\phi$ and $\epsilon$ follow from equation (6b). To solve for $\phi$ as a function of $\epsilon$, consider the completion of one word with all the optional extra letters completed. This set up gives $M = M_{max}$ per word (either \$0.20 or \$0.30), $E = 0$ and $S = \$0.90$, as $S$ is the monetary value of the letters completed in absence of pro-environmental effort. Thus, $\phi$ can be solved for as a function of $\epsilon$.

[6]The slope of the lines is given by $1/\phi$. The slope is calculated by taking subject earnings for a round, $Y_i = S_i - M_i$. Earnings sacrificed for pro-environmental effort, $M_i$, can be substituted for from equation (5), giving the equation for the relation between earnings and damages avoided, $Y_i = S_i - \frac{1}{\phi}E_{Mi}$.

[7]Only one session of 24 subjects was required for treatment group B as this group is not used to test any hypotheses on its own. Treatment group B boosts the numbers for the high damage to low damage combination from round 1 to 2 of treatment group A, and the high damage to high income combination from round 1 to round 3 combination of treatment group E.

Figure 2: Budget constraints by treatment, faced by a subject who can complete 126 letters in eight minutes.




### 3.2.2 Experimental procedures

The nine experimental sessions of 24 subjects each were conducted at the Monash Laboratory for Experimental Economics (MonLEE) at Monash University in Melbourne, Australia.[8] Current students of Monash University, registered in the MonLEE subject pool, were invited to attend a maximum of one session each. Sessions were conducted in June 2016 and May 2017. Student were invited to participate using ORSEE subject management system for the 2016 sessions (Greiner, 2015) and SONA for the 2017 sessions.[9] The study was named "A study of behaviors" so that the recruitment process was not biased towards students with an interest in environmental issues.

At the beginning of the session, the overview instructions were read, followed by the

[8]One session for treatment group A had only 23 subjects attend. Due to this lower number of subjects, the charitable donation for the rounds for that session was lowered by $14 and this was explained to subjects at the start of the session.

[9]The transition from ORSEE to SONA was managed such that no subject could participate in the study twice. See https://www.sona-systems.com/default.aspx for information about SONA (accessed 29 April 2017).

Table 2: Treatment groups by treatment order plus number of subjects in each group.

| Treatment group | Round 1 | Round 2 | Round 3 | Number of subjects |
|---|---|---|---|---|
| A | High damage | Low damage | Low effort | 47 |
| B | High damage | Low damage | High income | 24 |
| C | Low damage | High damage | Low effort | 48 |
| D | High damage | Choice | Low effort | 48 |
| E | High damage | Choice | High income | 48 |
| Total subjects | | | | 215 |

activity instructions (see Appendix). Next, subjects undertook a simple and incentivized quiz to ensure they understood the instructions. There were four questions, with each correct answer worth 25c. Subjects were informed by the software immediately after submitting their answer whether or not they were correct. If incorrect, the correct answer was given and explained.

The earnings and damages per word for each round were read aloud and displayed on the screen before each round to establish common knowledge that every subject had the same incentives for the round. After the three rounds, a survey was given to subjects – the variables used from the survey in the analysis are described in Section 4.1.

All activities and the survey were conducted using the z-Tree program (Fischbacher, 2007). After the survey, the experimenter announced the round that would be paid, including the total session-level payment to the charity. It was explained at the start of the activities that the charity payment would not be known until this point, and that proof of the donation would be provided via email in the days after the experiment had finished. Finally, each subject was paid in private. Each session lasted roughly one hour.

## 3.3 Hypotheses

### 3.3.1 Within subject hypotheses

The hypotheses are described here in relation to the theoretical model of Section 3.1.2. They are separated into within and between subject hypotheses. The first within subject hypothesis is as follows:

**H1** The behavioral rebound effect is positive.

The behavioral rebound effect is equivalent to the negative of the energy efficiency elasticity of pro-environmental effort, $-\eta_\epsilon(M)$. That this value is positive follows from its definition in equation (9), and the assumption given in equation (7). Specifically, an increase in energy efficiency, $\epsilon$, has a negative impact on the effectiveness of pro-environmental effort, $\phi$, by assumption and by experimental design. A decrease in $\phi$ reduces the benefit/cost ratio for pro-environmental effort, which I hypothesize will lead to a decrease in pro-environmental effort, $M$. This decrease in pro-environmental effort is consistent with assuming pro-environmental effort is undertaken both for pro-environmental preferences and beliefs about pro-social norms. Thus, hypothesis H1 is that $-\eta_\epsilon(M) > 0$.

Hypothesis H2 is related to the difference in the treatments where $\phi$ is varied but $\epsilon$ is the same (high damage compared with the low effort treatments). Thus, looking just at the effectiveness elasticity of pro-environmental effort, $\eta_\phi(M)$, and consistent with hypothesis H1, the second hypothesis is:

**H2** The effectiveness elasticity of pro-environmental effort, with no change in energy efficiency, is positive.

Finally, hypothesis H3 is related to the change in pro-environmental effort between the high damage and high income treatments. As the behavioral rebound effect is an elasticity, it is composed of an income effect and a substitution effect, which could operate in the same or opposite direction. Hence, I estimate the income elasticity of pro-environmental effort, call it $\eta_Y(M)$, to determine the magnitude and direction of the income effect for $\eta_\phi(M)$. The difference in pro-environmental effort between the high income treatment and the high damage treatment tests the direction of the income effect in this context. The Environmental Kuznet's Curve hypothesis suggests that pro-environmental preferences rise with income as environmental quality is a luxury good (Dinda, 2004). I thus hypothesize that:

**H3** There is a positive income elasticity of pro-environmental effort.

### 3.3.2 Between subject hypotheses

The first between subject hypothesis relates to the level of pro-environmental effort measured within a given treatment round. It is that:

**H4** Pro-environmental effort can be partially explained by demographics, environmental values and beliefs about social norms.

Given the literature outlined in the introduction and background sections, in Section 3.1.2 I argue that level of pro-environmental effort is driven by pro-environmental preferences and aversion to deviating from social norms. I measure beliefs about social norms directly and thus hypothesize that higher levels of beliefs about social norms of pro-environmental effort will drive higher levels of pro-environmental effort. Pro-environmental preferences will likely be formed through a combination of life experience and pro-environmental beliefs and values. There will be little variation in demographics in the data, given the relatively homogeneous subject pool, but measures of pro-environmental orientation along with observed pro-environmental behaviors are still likely to be highly heterogeneous, given findings of other studies (Hawcroft and Milfont, 2010; Sturm and Weimann, 2006). I describe the relevant variables used for this hypothesis in Section 4.1.

The next between subject hypothesis is that:

**H5** There is a moral licensing effect. Specifically, the drop in pro-environmental effort will be less when moving from the high to low damage treatments compared with the high to low damage choice treatments.

This hypothesis follows from the moral licensing literature, as previously described. It is tested by comparing the treatment groups who were given the low damage treatment exogenously to the treatment groups who were given a costless choice between high and low damages per word. The choice is costless, meaning that the only difference between the two treatments is that the choice itself is the only difference between the two treatments. There are no differences in earnings for that round between the subject groups. Hence, subjects may give themselves a moral license to put in less effort after choosing low damages compared with when they have low damages imposed exogenously. I test this hypothesis using a differences-in-differences approach, as stated in the hypothesis itself.

Finally, I look at whether there is heterogeneity in the moral licensing effect due to pro-environmental orientation of their attitudes, values and beliefs. Moral licensing occurs when an individual undertakes a moral action - in this case, choosing low damages over high damages for a round - and then give themselves a psychological license to undertake less moral behaviors after that point than they otherwise would. This effect thus hinges on the individual seeing the action they have undertaken as moral in the first place. Thus, I hypothesize that subjects with a higher pro-environmental orientation will see choosing low damages as more of a moral choice than those with a lower pro-environmental orientation, where pro-environmental orientation is a mix of values and beliefs about the environment. Hence, the final hypothesis is:

**H5a** The moral licensing effect is larger for those with a higher pro-environmental orientation.

# 4 Results

In the first part of this section I present the summary statistics while clearly defining the relevant variables. Next, I present the econometric analysis of the results. I finish the section by outlining each main result as it pertains to the relevant hypothesis.

## 4.1 Summary statistics

The outcome variable of interest in this experiment is proportion of pro-environmental effort. This variable is calculated as the proportion of the optional extra pro-environmental letters completed out of the total possible, for each individual in each round.[10] As this measure is robust to number of letters completed in a round, it is suitable to use to compare both within individuals (it allows for variation of letters completed within rounds, for example due to an error in one of the words) and between individuals (it allows for difference in overall effort and/or skill). As it is a proportion, it is a continuous variable on the unit interval. This variable is used as the dependent variable throughout the results section.

Proportion of pro-environmental effort by treatment is summarized in the top half of Table 3a. The treatment with the lowest effort is choice - chose low, which is the treatment where subjects could choose between high and low damages and includes just those who chose low damages (85 out of the 96 subject given this treatment). Proportion of pro-environmental effort in this treatment is 0.23, meaning less than one of the three optional damage reducing letters were completed per word. Next lowest treatment by pro-environmental effort is high income, followed by low damage, high damage, low effort and finally choice - chose high.

It is useful to note at this stage that no learning effect is observed; the mean total number of letters completed per round is almost identical over each round. The mean letters completed for all treatment groups in order of round are 125.7, 126.0 and 126.3, with no statistical difference detected ($0.72 > p > 0.88$, depending on the rounds compared). Mean letters completed by treatment group are similarly stable. Thus, the randomized

[10] The variable is calculated so that it takes into account that individuals who complete more optional extra letters will complete the roughly same number of letters in a round, but will complete fewer words, as the optional extra letters make the words longer. Consider an individual who completes 126 letters in a high damage treatment round. If they complete all the optional extra letters, they will complete 14 words ((126/9), which is $14 \times 3 = 42$ extra letters. Thus, the maximum number of extra letters they could complete at a rate of 126 letters per round is 42. Hence, if they complete 1 optional extra letter per word, they complete $126/7 = 18$ words, thus 18 optional extra letters and $18/42 = 0.43$ optional extra letters out of the total they could complete. This example demonstrates that if an individual completes 1 of 3 optional extra letters for each word they complete, rather than the proportion of optional extra letters being completed being 0.33, it is actually 0.43 as the calculation needs to take into account the total number of words they complete is falling as they complete more extra letters.

Table 3: Summary statistics.

(a)

| Statistic | N | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|
| **Proportion pro-environmental effort by treatment** | | | | | |
| High damage | 215 | 0.32 | 0.39 | 0 | 1 |
| Low damage | 119 | 0.29 | 0.40 | 0 | 1 |
| Choice – chose low | 85 | 0.23 | 0.35 | 0 | 1 |
| Choice – chose high | 11 | 0.42 | 0.45 | 0 | 1 |
| Low effort | 143 | 0.36 | 0.40 | 0 | 1 |
| High income | 72 | 0.25 | 0.36 | 0 | 1 |
| **Continuous covariates** | | | | | |
| Age | 215 | 21.81 | 3.77 | 17 | 48 |
| Norm belief | 215 | 1.11 | 0.96 | 0 | 3 |
| Letters high damage | 215 | 125.77 | 21.14 | 63 | 168 |
| Environmental behaviours | 215 | 3.64 | 0.44 | 2.62 | 4.93 |
| NEP scale | 215 | 3.72 | 0.47 | 2.60 | 4.87 |

(b)

| Statistic | N | % |
|---|---|---|
| **Gender** | | |
| Female | 110 | 51.2 |
| Male | 105 | 48.8 |
| **Subjective personal income** | | |
| Low | 182 | 84.7 |
| Medium | 33 | 15.3 |
| High | 0 | 0 |
| **Citizenship** | | |
| Australian | 48 | 22.3 |
| Not Australian | 167 | 77.7 |
| **Environmental organization** | | |
| Not member | 182 | 84.7 |
| Member | 33 | 15.3 |
| **Political party** | | |
| Not member | 206 | 95.8 |
| Member | 9 | 4.2 |
| **Voting preference** | | |
| Liberal | 36 | 16.7 |
| Labor | 27 | 12.6 |
| Greens | 19 | 8.8 |
| Other | 7 | 3.3 |
| Unsure | 126 | 58.6 |
| **Total** | 215 | 100 |

alphabet design from Benndorf et al. (2018) successfully prevented learning effects from affecting the results.

The other summary statistics in Table 3 are primarily the subject responses to the survey given at the end of each experimental session. The bottom half of Table 3a gives the summary statistics for the continuous covariates. Most subjects are close to the mean age of 21.8, as expected from a standard student subject pool. The norm belief variable gives the subject response to the question of what they believe to be the average number of optional extra letters in round 1 of other subjects in their session. Subjects could only answer in whole numbers between 0 and 3, and on average guessed the correct number, 1. The letters high damage variable is the number of letters completed by subjects in the high damage treatment. Environmental behaviors is a measure of stated frequency of undertaking pro-environmental behaviors within the last year, between 1 (never) and 5 (always). The measure is produced by averaging the response to all the environmental behavior questions included in the survey, for which a Likert scale was employed. Finally, the New Ecological Paradigm (NEP) scale is a measure of pro-environmental orientation of attitudes and beliefs (see Appendix for questions used for these latter two variables). This is also a variable utilizing a Likert scale from 1 to 5, depending on answers to a standard 15 question survey on environmental values and attitudes, where a higher number denotes a stronger pro-environmental orientation (Dunlap et al., 2000). The mean value of 3.7 falls within 0.1 of the mean value recorded for two 15 question NEP surveys undertaken in Australia in roughly the last decade (Hawcroft and Milfont, 2010).

Table 3b shows the discrete variables. First, the gender balance is very even, with 51% of subjects being female. Subjective personal income level stated by subjects is mostly low (85%), with the rest being medium. This pattern is not unexpected with a student subject pool. Subjective variables such as level of income often prove to be informative explanatory variables (Bertrand and Mullainathan, 2001). Next, the sample has a large number of subjects who are not Australian citizens (78%), which simply reflects the subject pool at the MonLEE lab. There is no particular reason to believe using a largely non-Australian subject pool would affect the testing of the hypotheses, but with collecting data on citizenship I can control for this variable. The next two variables are subject responses to whether they have ever been a member of a environmental organization or political party, to indicate political engagement, particularly concerning environmental issues. Not many subjects report being or having been a member of either (15% and 4% respectively). Finally, the voting preference question asked subjects which political party they would give their first preference to if voting on the day of the survey, and regardless of their Australian citizenship status. Of note to the research question is 9% stating they would vote for the Greens Party. A majority stated they were unsure at 59%, which is unsurprising given the large number of non-citizens.

## 4.2 Econometric analysis

Here I outline the main econometric approach and introduce the overall results. In Sections 4.3 and 4.4 I discuss in detail the results as they pertain to each hypothesis.

The within subject hypotheses depend on the difference in pro-environmental effort between particular treatments; the difference from high to low damage treatments, the difference between high damage to low effort treatments and the difference between high damage and high income treatments. Thus, I test whether the differences in pro-environmental effort between these treatments are statistically significant and in the direction consistent with the first three hypotheses. Table 4 tests these differences using the non-parametric Mann-Whitney U test. The first row in Table 4 tests the difference between the high damage and low damage treatments, specifically testing whether the proportion of pro-environmental effort in the low damage round minus the high damage round is negative. This test is done for all 119 subjects who received both treatments. The result is negative, with a p-value of 0.001. The second row in Table 4 tests the difference between the high damage and low effort treatments, where the low effort treatment has the same damage level as the high damage treatment per word, but only requires two optional additional letters to be completed to reduce damages to 0 for each word, rather than 3. The final row tests tests the difference between the low and high income treatments, specifically testing whether pro-environmental effort in the high income round minus the high damage round

Table 4: Testing for differences between treatments in the direction relevant to the within subject hypotheses, using the non-parametric paired Wilcoxon signed-rank test.

| | Negative | Positive | N |
|---|---|---|---|
| **Difference** | | | |
| High to low damage | 0.001*** | | 119 |
| High to low effort | | 0.185 | 143 |
| Low to high income | | 0.738 | 72 |

Note: *p<0.1; **p<0.05; ***p<0.01.

is positive. Neither of these two values are found to be statistically positive.[11]

Estimated treatment effects are shown in Table 5. The first two columns show model (1), which regresses dummy variables for each treatment round, relative to high damage, and each treatment group, relative to treatment group A, on the proportion of pro-environmental effort. A Tobit model is used given the dependent variable is subject to corner solutions (Wooldridge, 2010). I discuss this choice of model in more detail shortly. The left column for model (1) shows the estimated coefficients, whereas the right column shows the average marginal effects (AME). Average marginal effects are of more interest in this paper as they show the average effects of the treatments for the subject pool on the proportion of pro-environmental effort.[12] Each model in Table 5 uses all three observations from all 215 subjects, thus standard errors are clustered at the subject level.

Focusing on the low damage and low effort coefficients in model (1), Table 5, the former is negative and statistically significant and the latter is positive but not statistically significant, consistent with the results in Table 4. The treatment group controls are important to include to remove any differences between the average effort levels of the treatment groups, random or otherwise, as is standard with differences-in-differences

[11]It could be argued that total effort (total number of letters completed within a round) could also be increased from low to high income treatments, given the higher reward per letter in the high income treatment. I test for whether there is an increase in total effort between these treatments using a one-sided, paired Wilcoxon signed-rank test, and find it is not statistically significant ($p = 0.161$), therefore it does not affect the analysis in any significant way.

[12]The Tobit coefficients can be interpreted as the estimated marginal effect of each variable if there were no corner solutions, whereas the AMEs provide the mean marginal effect of each variable for the subjects in the sample, taking into account that some subjects are at the corner solutions. Thus, the AMEs are more informative as they can be interpreted as the marginal change from the treatments in the expected proportion of pro-environmental effort at the population level. Wooldridge (2010) refers to these as average partial effects, or average treatment effects when referring to dummy variables.

Table 5: Tobit models testing treatment effects

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Proportion pro-environmental effort | | | |
| | (1) | | (2) | |
| | Coefs | AMEs | Coefs | AMEs |
| Low damage | -0.1178** | -0.0461*** | -0.1572** | -0.0609** |
| | (0.0465) | (0.0178) | (0.0640) | (0.0239) |
| Low effort | 0.0081 | 0.0033 | -0.0478 | -0.0189 |
| | (0.0361) | (0.0145) | (0.0638) | (0.0249) |
| Income effect | -0.0292 | -0.0116 | -0.0168 | -0.0067 |
| | (0.0367) | (0.0145) | (0.0348) | (0.0139) |
| Choice | 0.1442 | 0.0589 | 0.1935 | 0.0795 |
| | (0.1494) | (0.0620) | (0.2146) | (0.0896) |
| Chose low | -0.0609 | -0.0243 | 0.2014 | 0.0806 |
| | (0.2899) | (0.1150) | (0.3023) | (0.1204) |
| Choice*Chose low | -0.3409** | -0.1254** | -0.4881** | -0.1715*** |
| | (0.1543) | (0.0510) | (0.2147) | (0.0625) |
| TG B | -0.3814* | -0.1382* | -0.3971* | -0.1432** |
| | (0.2249) | (0.0714) | (0.2270) | (0.0709) |
| TG C | -0.0277 | -0.0110 | -0.0587 | -0.0232 |
| | (0.1872) | (0.0741) | (0.1889) | (0.0739) |
| TG D | 0.0528 | 0.0213 | -0.1891 | -0.0727 |
| | (0.2964) | (0.1204) | (0.3103) | (0.1141) |
| TG E | -0.1355 | -0.0531 | 0.3227 | 0.1320 |
| | (0.3365) | (0.1287) | (0.5824) | (0.2395) |
| Low damage*TG B | | | -0.0186 | -0.0074 |
| | | | (0.1015) | (0.0402) |
| Low damage*TG C | | | 0.0668 | 0.0271 |
| | | | (0.1111) | (0.0457) |
| Low effort*TG C | | | 0.0276 | 0.0111 |
| | | | (0.0973) | (0.0394) |
| Low effort*TG D | | | 0.0943 | 0.0385 |
| | | | (0.0865) | (0.0358) |
| Inc. effect*TG E | | | 0.0294 | 0.0118 |
| | | | (0.0596) | (0.0242) |
| Choice*TG E | | | -0.1092 | -0.0425 |
| | | | (0.2234) | (0.0840) |
| Chose low*TG E | | | -0.8444 | -0.2729* |
| | | | (0.6576) | (0.1502) |
| Ch.*Ch. low*TG E | | | 0.3444 | 0.1453 |
| | | | (0.2318) | (0.0989) |
| Constant | 0.2008 | | 0.2335* | |
| | (0.1267) | | (0.1243) | |
| $\hat{\sigma}$ | 0.6719*** | | 0.6650*** | |
| | (0.1060) | | (0.1054) | |
| N | | 215 | | 215 |
| P-value | | 0.0010 | | 0.0007 |
| Pseudo r-squared | | 0.0148 | | 0.0190 |
| Pseudo log-lik. | | -624.82 | | -622.18 |

Notes: TG abbreviates "Treatment group". Standard errors are clustered at the subject level and in parentheses. The delta-method is used to calculate standard errors for average marginal effects (AMEs). *p<0.1; **p<0.05; ***p<0.01.

models.[13]

Model (2) in Table 5 includes all relevant interactions between treatments and treatment groups, given which treatment groups received which treatment. Thus, these interactions allow me to test whether the treatment effects are also affected by any differences between the treatment groups. There are no significant coefficients on the treatment-treatment group interactions. Including these interactions does increase the absolute size of the low damage coefficient, and changes the low effort coefficient from positive to negative, but it remains not statistically significant. Another important test from this model is for any order effects as treatment groups A and B saw the high damage round first and low damage second, whereas treatment group C saw the rounds in the opposite order. The instructions and practice round were consistent with the first round that subjects were given to fully control for any order effects. There is no statistical significance on the coefficient of low damage*TG C, thus there is no evidence that order effects are a significant driver of the results. Given the lack of significance on any of the treatment and treatment group interaction coefficients, and using the AIC and BIC criteria, model (1) is chosen as the preferred model over model (2) for the analysis of the hypotheses.

At this point it is useful to briefly discuss the choice of the Tobit model. The Tobit model is used in this analysis as it is a corner solution model (Wooldridge, 2010) and a large number of the dependent observations are corner solutions on the unit interval (47% at 0 and 13% at 1 for proportion of pro-environmental effort). The Tobit model accounts for the corner solutions that subjects have arrived at by modeling a probability that a subject has chosen 0 or 1. Average marginal effects in the Tobit thus account for the non-linearities in the dependent variable, which are not accounted for by OLS (Wooldridge, 2010). Nevertheless, OLS estimates are provided in the Appendix as a robustness check.

[13]The different systems used for recruitment for some sessions, as noted in Section 3.2.2, may lead to some minor differences between the treatment groups; otherwise there is no other systematic difference between the sessions. Specifically ORSEE was used for treatment groups A, C and D, whereas SONA was used for treatment groups B and E. SONA allows all eligible subjects in the subject pool to sign up to any experimental session, whereas ORSEE only allows a random subset of subjects who receive an email invitation to sign up to a session. Treatment group B does have a statistically significantly lower effort level than treatment group A, thus it is important to conduct the analysis with the treatment group controls. The recruitment differences are not expected to affect the results regarding the hypotheses; this assumption is confirmed by the results of model (2).

Table 6: Estimated elasticities.

| Parameters | High to low damages | High to low dam., moral licensing | High to low effort | Income effect |
|---|---|---|---|---|
| $\Delta M/M$ (est.) | $-0.1593^{**}$ | $-0.4910^{**}$ | $-0.2615^{***}$ | $-0.0388$ |
| $\Delta\epsilon/\epsilon$ | 0.5 | 0.5 | | |
| $\Delta\phi/\phi$ | $-0.3333$ | $-0.3333$ | 0.3333 | |
| $\Delta Y/Y$ | | | | 0.3333 |
| $-\eta_\epsilon(M)$ (est.) | **0.32** ($\pm0.26$) | **0.98** ($\pm0.95$) | | |
| $\eta_\phi(M)$ (est.) | **0.48** ($\pm0.39$) | **1.47** ($\pm1.42$) | **−0.78** ($\pm0.31$) | |
| $\eta_Y(M)$ (est.) | | | | **−0.12** ($\pm0.29$) |

Notes: Elasticities are estimated from model (1) of Table 5, though high to low effort is adjusted to account for lower absolute effort being required in the low effort treatment. The parameters in the upper section of the table are chosen such that the elasticities of interest can be calculated; for example $\eta_\epsilon(M)$ is estimated by dividing $\Delta M/M$ by $\Delta\epsilon/\epsilon$, as per the standard definition of an elasticity. Standard errors are calculated using the delta-method. Brackets contain 95% confidence intervals. *p<0.1; **p<0.05; ***p<0.01.

There are no unexpected or particularly significant differences between the estimates from the Tobit and OLS models.

I also undertake a rough test for misspecification of the Tobit model. The test is to compare Tobit coefficients, divided by the estimated variance, with Probit coefficients for models estimated on dummies for 0 and greater than 0, and less than 1 and 1 (Wooldridge, 2010). The results of this exercise indicate no issues of misspecification. In terms of alternative models, the Cragg hurdle model (Wooldridge, 2010) was experimented with and provided no notable differences with the Tobit. Thus, the Tobit was favored as the more parsimonious model.

Table 6 shows the estimated elasticities from the preferred model, model (1) of Table 5. The middle section of the table shows the values used to calculated the elasticities in the bottom section. The first row in the bottom section shows the negative of the energy efficiency elasticity of pro-environmental effort, $-\eta_\epsilon(M)$, which I defined in Section 3.1.2 as the behavioral rebound effect. This is positive and estimated as 0.32 for high to low damages and 0.98 for high to low damages with moral licensing. The numbers in brackets next to the elasticity estimate give the 95% confidence intervals.

The next elasticity in Table 6 is $\eta_\phi(M)$, the effectiveness elasticity of pro-environmental effort, which relates only to how much damage reduction is achieved by an additional unit of pro-environmental effort. Given the effectiveness of pro-environmental effort, $\phi$, is a function of energy efficiency, $\epsilon$, for the change from high to low damages, $\eta_\phi(M)$ is closely related to $-\eta_\epsilon(M)$ for the first two columns. However, the high to low effort treatment change is constructed so that $\phi$ increases while $\epsilon$ stays the same, hence only $\eta_\phi(M)$ can be estimated for this difference in effort between treatments. For high to low effort, $\eta_\phi(M)$ is estimated to be -0.78. Given there are only two optional additional letters in the low effort treatment instead of three, the measure of the proportion of pro-environmental effort is adjusted to account for the fact that completing one extra letter for every word is a 0.43 proportion of pro-environmental effort for the high damage treatment, but a 0.57 proportion of pro-environmental effort for the low effort treatment. This elasticity is thus calculated after adjusting proportion of pro-environmental effort to be an equivalent of absolute level of effort between treatments, by multiplying proportion of pro-environmental effort for the low effort treatment by 3/4. Finally, the income elasticity of pro-environmental effort $\eta_Y(M)$, is given for the high income treatment relative to the high damage treatment. It is small and not statistically significant, as per the estimate in model (1), Table 5.

Table 7 shows the proportion of pro-environmental effort, regressed on the responses to the survey, using a Tobit model. Only the high damage treatment is included for treatment groups that saw the high damage treatment for the first round. Only these observations are included as some of the survey questions were given in regards to the first round, thus it is important only to compare groups receiving the same treatment in the first round. Results on other treatments and rounds are consistent with these results.

Model (1) in Table 7 regresses just the norm belief variable on pro-environmental effort, and shows this variable is a strong driver of pro-environmental effort. The coefficient is highly significant and positive, with the belief that one extra letter is completed by others on average being associated with a 0.21 increase in proportion of pro-environmental effort. Model (2) shows the range of controls available being regressed on pro-environmental effort, without the norm belief variable. In this model, gender and reported levels of pro-environmental behaviors are significant explanatory variables. Model (3) includes all relevant variables. There is little change to the coefficient on the norm belief variable

Table 7: Tobit models of proportion pro-environmental effort in high damage treatment, treatment groups A, B, D, E.

| | *Dependent variable:* Proportion pro-environmental effort | | | | | |
|---|---|---|---|---|---|---|
| | (1) | | (2) | | (3) | |
| | Coefs | AMEs | Coefs | AMEs | Coefs | AMEs |
| Norm belief | 0.4526*** | 0.2098*** | | | 0.4213*** | 0.1967*** |
| | (0.0527) | (0.0152) | | | (0.0525) | (0.0174) |
| Letters | | | -0.0031 | -0.0015 | 0.0003 | 0.0001 |
| | | | (0.0027) | (0.0013) | (0.0022) | (0.0010) |
| Age | | | 0.0196 | 0.0092 | 0.0121 | 0.0057 |
| | | | (0.0173) | (0.0081) | (0.0135) | (0.0063) |
| Female | | | 0.3507*** | 0.1644*** | 0.1404 | 0.0661 |
| | | | (0.1134) | (0.0508) | (0.0898) | (0.0423) |
| Low income | | | -0.1699 | -0.0828 | -0.1537 | -0.0745 |
| | | | (0.1523) | (0.0764) | (0.1189) | (0.0596) |
| Australian | | | 0.0387 | 0.0183 | 0.0088 | 0.0041 |
| | | | (0.1413) | (0.0672) | (0.1112) | (0.0521) |
| Env. behav. | | | 0.2733** | 0.1285** | 0.1378 | 0.0643 |
| | | | (0.1234) | (0.0566) | (0.0962) | (0.0447) |
| NEP scale | | | -0.0344 | -0.0162 | 0.0164 | 0.0076 |
| | | | (0.1275) | (0.0599) | (0.1003) | (0.0468) |
| Env. org. | | | 0.1841 | 0.0904 | 0.1110 | 0.0537 |
| | | | (0.1650) | (0.0841) | (0.1289) | (0.0643) |
| Political party | | | 0.4913* | 0.2529 | 0.2505 | 0.1263 |
| | | | (0.2938) | (0.1542) | (0.2258) | (0.1206) |
| Vote Labor | | | -0.3292 | -0.1415* | -0.2900* | -0.1249* |
| | | | (0.2096) | (0.0795) | (0.1660) | (0.0642) |
| Vote Greens | | | 0.2305 | 0.1142 | 0.2473 | 0.1233 |
| | | | (0.2409) | (0.1239) | (0.1879) | (0.0980) |
| Vote other | | | -0.0722 | -0.0331 | -0.2580 | -0.1081 |
| | | | (0.3539) | (0.1583) | (0.2818) | (0.1038) |
| Vote unsure | | | 0.0843 | 0.0394 | -0.0474 | -0.0221 |
| | | | (0.1600) | (0.0742) | (0.1268) | (0.0594) |
| Constant | -0.3331*** | | -0.8751 | | -1.0714 | |
| | (0.0828) | | (0.8307) | | (0.6590) | |
| $\hat{\sigma}$ | 0.2663*** | | 0.3967*** | | 0.2303*** | |
| | (0.0469) | | (0.0717) | | (0.0406) | |
| N | | 167 | | 167 | | 167 |
| Observations | | 167 | | 167 | | 167 |
| P-value | | 0.0000 | | 0.0099 | | 0.0000 |
| Pseudo r-squared | | 0.2441 | | 0.0873 | | 0.2989 |
| Log-likelihood | | -120.00 | | -144.91 | | -111.31 |

Notes: The delta-method is used to calculate standard errors for average marginal effects (AMEs). The "vote" dummy variables are relative to Vote Liberal. *p<0.1; **p<0.05; ***p<0.01.

compared with model (1), whereas the controls change in size and statistical significance. Treatment group controls were tested, but were not statistically significant and did not change the overall results. Their inclusion also worsened model fit using the AIC and BIC criteria, and thus have not been included in Table 7.

## 4.3 Within subject hypotheses

### 4.3.1 The behavioral rebound effect

I now discuss the results as they pertain to hypothesis H1, that there is a positive behavioral rebound effect. If the difference in pro-environmental effort of the low damage treatment minus the high damage treatment is negative, then the behavioral rebound effect will be positive. Table 4 finds that this difference between treatments is negative at the 1% level. This result is also supported by the Tobit models estimated in Table 5, with the negative and statistically significant coefficient on the low damage treatment dummy, which is relative to the high damage treatment. Thus, the first main result, corresponding to hypothesis H1, is:

**Result 1** There is a positive behavioral rebound effect.

Thus, I fail to reject hypothesis H1. The behavioral rebound effect for high to low damages is estimated to be 0.32, as shown in Table 6.

### 4.3.2 Effectiveness elasticity of pro-environmental effort

Hypothesis H2 is that the effectiveness elasticity of pro-environmental effort, when there is no associated change in energy efficiency, is positive. This hypothesis is tested through estimating the difference between the high damage treatment and the low effort treatment. There is no positive effect found in the non-parametric testing in Table 4, nor on the low effort coefficient in Table 5. The results presented in Table 6 for the estimated elasticity controls for the difference in absolute pro-environmental effort, rather than just proportion of pro-environmental, given low effort requires just two optional extra letters to be completed per word, rather than three. The following result is found:

**Result 2** The effectiveness elasticity of pro-environmental effort associated with an increase in the effectiveness of pro-environmental effort with no change in energy efficiency is estimated to be -0.78.

Thus, I reject hypothesis H2. In fact, the increase in effectiveness of pro-environmental effort, through an increase in $\phi$, implemented by lowering the effort required to lower damage, is met with a strong reduction in pro-environmental effort. Taking into account the confidence interval, I cannot rule out an elasticity of -1. This finding is particularly interesting given, as also shown in Table 6, the effectiveness elasticity of pro-environmental effort is positive when the change in effectiveness of pro-environmental effort, $\phi$, is due to a change in energy efficiency, $\epsilon$.

#### 4.3.3 Income effect

In terms of hypothesis H3, the non-parametric Mann-Whitney U test for income effect in Table 4 shows no evidence for an income effect forming a part of the behavioral rebound effect. There is also no statistically significant income effect estimated in Table 5. These results are a rejection of hypothesis H3.

The change in income between the high damage and high income treatments is one third; perhaps this is not a large enough change on the income side to have an effect on pro-environmental effort. Another test available is whether subject income level has any effect on their pro-environmental effort, though given subjects are students there is little variation in their personal income. The results in Table 7 again present no evidence for personal income having a positive effect on pro-environmental effort. Also included in Table 7 is whether number of letters a subject completes in a round (which also impacts income earned for the experiment) has an impact on pro-environmental effort. The coefficient on that variable is not significantly different from 0. Thus, the following result is found:

**Result 3** There is no evidence of a positive income effect on pro-environmental effort.

## 4.4 Between subject hypotheses

### 4.4.1 Pro-environmental effort and observables

I first discuss hypothesis H4, which relates social norms, demographics and environmental values to level of pro-environmental effort. Table 7 shows the main results for this hypothesis, with model (1) showing a strong link between pro-environmental effort and beliefs about social norms. Model (2) of Table 7 shows that females are more inclined to put in pro-environmental effort, as are people who report more pro-environmental behaviors in their daily life. Interestingly, pro-environmental orientation, as measured by the NEP scale, does not statistically predict pro-environmental effort when controlling for other variables. Australian citizen is another variable that could be hypothesized to positively predict pro-environmental effort, given the tree planting charity plants indigenous trees within the state of Victoria, but it is not statistically significant. However, model (3) includes all relevant variables and supports the overall finding:

**Result 4** The strongest driver of pro-environmental effort is beliefs about social norms.

Including both social norms and other variables in model (3) leads to the size and significance of the female and environmental behavior coefficients dropping away relative to model (2), which does not include beliefs about social norms. Norm beliefs are positively correlated both with being female and environmental behaviors and model (3) suggests that norm belief itself is the most powerful driver of pro-environmental behaviors.

One potential criticism of the norm belief variable is that subjects just chose the number closest to their level of effort, given there is no incentive to consider the question more deeply. To address this criticism, treatment groups B and E were incentivized for the norm belief question, being told they would earn $1 if they chose the average number of optional extra letters completed closest to the actual level in their session. There is no statistical difference between subjects incentivized for this question and subjects not incentivized ($p = 0.73$), thus the incentivized and non-incentivized treatment groups are pooled for Table 7.

#### 4.4.2 Moral licensing

I now address the final set of coefficients not yet addressed in Table 5. These coefficients are the choice treatment dummy, chose low and the interaction between choice and chose low.[14] Formally, to be consistent with hypothesis H5, the choice*chose low coefficient must be less than the low damage coefficient. Hence, I conduct a one-sided $t$ test that the coefficient on choice*chose low $<$ low damage. For both models, the difference is significant at the 10% level, hence:

**Result 5** There is evidence consistent with moral licensing.

Thus, I fail to reject hypothesis H5. The 10% statistical significance level of the moral licensing behavior is also consistent with the meta-study on moral licensing, Blanken et al. (2015). They find that moral licensing has a small to medium effect size, and thus for a 5% significance level with statistical power of 80% the study would need 165 subjects each in a control and moral licensing treatment group. This sample size is within the norm for a laboratory experiment, given the cost of the method, but evidence of moral licensing at the 10% level with the sample size in this study is consistent with Blanken et al.'s (2015) estimated effect size of moral licensing.

The results in Table 6 show that the behavioral rebound effect is estimated to be roughly three times the size under moral licensing, compared with exogenous technological change. At 98%, this is a large rebound, pushing costless endogenous technological change roughly at the level of backfire. However, this estimated behavioral rebound effect does have a large confidence interval associated with it, so neither a much lower moral licensing effect nor backfire can be ruled out.

[14]The baseline pro-environmental effort of subjects who chose high damages in the choice treatment are accounted for in the treatment group dummies for groups D and E.

Finally, I test hypothesis H5a, that the moral licensing effect is larger for those with higher pro-environmental orientation. To test this hypothesis I re-estimate model (1) from Table 5 after separating the sample into two groups - those with a higher than the median pro-environmental orientation, according to the NEP measure, and those with less than or equal to the median pro-environmental orientation. The results are shown in Table 8.

Model (1) in Table 8 shows the results for those subjects with a equal to or stronger than median pro-environmental orientation. I conduct a one-sided $t$ test that the choice*chose low interaction coefficient is less than the low damage coefficient. This is significant at the 5% level. Model (2) thus shows the results for those subjects with a less than median pro-environmental orientation. Conducting the same one-sided $t$ test, the results are not statistically significant (with a p-value of 0.38). Hence:

**Result 5a** The moral licensing effect is larger for those with a higher pro-environmental orientation and is statistically insignificant for those with a lower pro-environmental orientation.

Therefore, I fail to reject hypothesis H5a. The results in Table 8 show a further interesting pattern: the other treatment effects are also different between the groups with low and high pro-environmental orientation. The low damage treatment has no statistical significance for those with high pro-environmental orientation, while it is larger for those with low pro-environmental orientation. This result suggests that the positive behavioral rebound effect is driven by those with lower pro-environmental orientation, as their pro-environmental effort is more sensitive to changing incentives. Additionally, there is a statistically significant negative income effect for those with high pro-environmental orientation. This latter result is something of a puzzle, though the effect size is not large.

To test whether Result 5a is driven by propensity to undertake pro-environmental effort, which seems to be driven mostly by beliefs about social norms, I conduct the same exercise from Table 8 in Table A.1 in the Appendix using the pro-environmental behaviors

Table 8: Tobits testing treatment effects, separated by NEP measure.

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Proportion pro-environmental effort | | | |
| | NEP > median | | NEP $\leq$ median | |
| | (1) | | (2) | |
| | Coefs | AMEs | Coefs | AMEs |
| Low damage | -0.0588 | -0.0279 | -0.1883*** | -0.0589*** |
| | (0.0616) | (0.0288) | (0.0693) | (0.0216) |
| Low effort | 0.0609 | 0.0296 | -0.0462 | -0.0148 |
| | (0.0398) | (0.0195) | (0.0684) | (0.0216) |
| Income effect | -0.0818** | -0.0385** | 0.0580 | 0.0189 |
| | (0.0363) | (0.0161) | (0.0720) | (0.0235) |
| Choice | 0.3877 | 0.1907 | -0.0187 | -0.0060 |
| | (0.2665) | (0.1286) | (0.0682) | (0.0219) |
| Chose low | 0.5063 | 0.2314* | -0.7383* | -0.2177** |
| | (0.3121) | (0.1255) | (0.4434) | (0.1109) |
| Choice*Chose low | -0.5440** | -0.2208*** | -0.2345* | -0.0720* |
| | (0.2674) | (0.0844) | (0.1259) | (0.0368) |
| TG B | -0.6712** | -0.2597*** | -0.1088 | -0.0343 |
| | (0.2951) | (0.0803) | (0.3461) | (0.1065) |
| TG C | -0.1537 | -0.0715 | -0.0162 | -0.0052 |
| | (0.2305) | (0.1030) | (0.3373) | (0.1082) |
| TG D | -0.7826** | -0.3016*** | 0.9967** | 0.3458** |
| | (0.3480) | (0.0919) | (0.4724) | (0.1453) |
| TG E | -0.6620* | -0.2593** | 0.2905 | 0.0957 |
| | (0.3929) | (0.1113) | (0.4896) | (0.1615) |
| Constant | 0.4244** | | 0.0177 | |
| | (0.1800) | | (0.1921) | |
| $\hat{\sigma}$ | 0.4348*** | | 0.9619*** | |
| | (0.0869) | | (0.2305) | |
| N | | 102 | | 113 |
| Observations | | 306 | | 339 |
| P-value | | 0.0099 | | 0.0079 |
| Pseudo r-squared | | 0.0467 | | 0.0375 |
| Pseudo loglik | | -279.62 | | -320.02 |

Notes: TG abbreviates "Treatment group". Standard errors are clustered at the subject level and in parentheses. The delta-method is used to calculate standard errors for average marginal effects (AMEs). *p<0.1; **p<0.05; ***p<0.01.

variable.[15] The pattern is not the same; neither the group with low nor the group with high levels of reported environmental behaviors has a statistically significantly significant moral licensing effect.

# 5 Discussion

The results presented above provide evidence for a positive behavioral rebound effect, and a negative effectiveness elasticity of pro-environmental effort when the change effectiveness of pro-environmental effort occurs without a change in energy efficiency. The lack of a statistically significant income effect on pro-environmental effort suggests that these findings are mostly driven by the substitution effect between private earnings and reducing environmental damages. The estimated size of the behavioral rebound effect is 32%, which is a similar size to the average direct rebound effect measured in the field (Gillingham et al., 2016; Sorrell et al., 2009). Thus, the behavioral rebound effect is shown to be significant, which suggests a need for further work into augmenting models of the rebound effect to include social norms and pro-environmental preferences. Indeed, the power of social norms in influencing pro-environmental behavior is highlighted in this study, as it has been in previous research (Allcott, 2011).

The strength that a laboratory experiment brings to this research is the ability to cleanly identify effects that might be difficult or impossible to identify in the field, namely pure pro-environmental behaviors. In the field, behaviors that are *prima facie* pro-environmental may in reality be undertaken for a range of other benefits that they might bring the individual as well as the environmental benefits they provide. The task of identifying pure pro-environmental behaviors is not helped by the fact that technologies are imperfect substitutes. A car and a bicycle both provide a transport service, but with vastly different associated attributes such as comfort, speed, fitness benefits and environmental damages. While co-benefits of pro-environmental behaviors are important, such as fitness and enjoyment, this novel experiment provides strong evidence that individuals

[15] It is difficult to conduct the equivalent exercise using the norm belief variable, as most respondents chose 1 and as a discrete variable it does not allow easy separation of the subjects to roughly two equal sized groups of high and low norm beliefs.

respond not only to private incentives that change with changing technology, but also incentives to put in effort for the environment.

The laboratory setting also has the desirable feature of allowing a clean distinction to be drawn between exogenous and endogenous technological change in order to test for moral licensing. Thus, the evidence in favor of moral licensing presented here is compelling, given it is demonstrated with a costless and essentially irrelevant choice and on a relatively small sample size with which to investigate moral licensing.

Perhaps most important to the literature on moral licensing is Result 5a. It shows an even larger and more significant moral licensing effect on the more pro-environmentally orientated half of the sample, where the 102 subjects in the high pro-environmental orientation subsample represent less than a third of the sample size recommended by Blanken et al. (2015) to measure moral licensing. Furthermore, this finding does seem to be related solely to underlying pro-environmental orientation and not to revealed pro-environmental effort, as shown when comparing Table 8, which uses the NEP scale to separate the models, with Table A.1, which uses reported level of pro-environmental behavior. It is a particularly interesting finding given that it is reported pro-environmental behavior that has some predictive power on the underlying level of pro-environmental effort in a given round, unlike the NEP scale, as shown in Table 7.

Nevertheless, the unique environment created in the laboratory also requires some caveats on the estimates of the elasticities. One main limitation in this study is the salience of environmental damage being much higher than the real world, given clear environmental damages with real time feedback to subjects. High salience is likely to encourage a higher level of pro-environmental behaviour than would be observed outside of the laboratory (Schubert, 2017).

Another potential limitation is that subjects are aware that they are being observed. While the data collected are anonymized, this aspect of the laboratory environment may still influence subjects to act more morally than they would in a private setting (Levitt and List, 2007). This limitation may mean that the overall level of pro-environmental effort is higher in the laboratory than in the field, but the size of the behavioral rebound effect may be smaller if the result of this attribute is also a smaller behavioral response to changes in the level of environmental damages. The same reasoning applies to moral licensing; it may also be smaller in the laboratory than in the field. The countervailing force to this is the fact that the behavioral rebound effect and moral licensing may be stronger given the higher salience of pro-environmental behaviors in the laboratory. With these limitations in mind, the laboratory environment still provides evidence that the behavioral rebound effect is significant and important and can help guide further research in the field. Additionally, all subjects face the same conditions and therefore the experiment is internally consistent for the purpose of testing the hypotheses.

Future research building on the novel experimental design used in this paper can add in more aspects of the rebound effect to help build a better picture of the relative importance of the behavioral rebound effect in the overall rebound effect in energy use. Another aspect that could be tested in future is the importance of real time feedback on environmental damage along with the effect of having uncertain environmental damages. It would also be worth investigating the power of social norms further. Areas to investigate include how information or priming about social norms at the beginning of the experiment might affect pro-environmental effort, rather than just asking subjects about their beliefs about social norms in the post-experiment survey.

More generally, this paper provides an impetus for more research to determine the importance in the field of the behavioral rebound effect and moral licensing. Careful thought must be put into developing theory and research that allows the identification of the behavioral rebound effect in the field. A welfare analysis of the rebound effect including the behavioral rebound effect would be useful to help analyze policy interventions. Aspects of the behavioral rebound effect that could be investigated in the field include looking for more evidence of moral licensing and how it operates over the short and long term after the purchase of a durable good. Finally, it would be worthwhile testing policy interventions to encourage pro-environmental behaviors within the context of ongoing technological change, where these policies are likely to be welfare enhancing.

# 6 Conclusion

In this paper I present a novel laboratory experiment, which provides both the ability to measure the level of pro-environmental effort, given a private cost to that effort, and how that effort changes with changing incentives. I find a behavioral rebound effect of around 32% associated with an increase in energy efficiency, which suggests that changes in pro-environmental effort contribute to the overall rebound effect in energy use. The results also show that technological changes that make pro-environmental effort easier or more effective, without changing energy efficiency, are unlikely to increase pro-environmental effort when accounting just for the environmental benefits. Additionally, I demonstrate the importance of beliefs about social norms for explaining level of pro-environmental effort. Finally, moral licensing increases the size of the behavioral rebound effect when technology change is endogenous, particularly for those individuals with a stronger pro-environmental orientation of their attitudes and beliefs.

It is worth considering a couple of examples to close, given the diversity of potential applications for the results of this paper. First, I return to the transport example for a final time. Those purchasing efficient cars for environmental reasons are subject to the direct rebound effect, the behavioral rebound effect and moral licensing. While these effects are unlikely to lead to backfire in the field, given most empirical evidence to date (Gillingham et al., 2016), the latter two effects could perhaps be reduced by information provision to continue to encourage pro-environmental behaviors, with a focus on social norms. Adding an electric assist to bicycles on the other hand is unlikely to encourage more cycling on environmental grounds, though it could still encourage more cycling if it enhances other benefits of cycling, such as enjoyment.

Another example is reducing greenhouse gas emissions from power generation. The purchase of rooftop solar panels may be subject to moral licensing, whereas building solar farms to reduce the carbon emissions of grid electricity is essentially exogenous from a consumer point of view, and therefore is unlikely to be subject to moral licensing. However, consumers choosing renewable energy options for their power provider are still potentially subject to moral licensing. Thus, policies to increase renewable energy in the power grid across all providers, along with continued energy conservation programs, may be more effective in reducing carbon emissions than subsidizing rooftop solar or relying on consumer demand for renewable power.

This paper provides strong evidence that it is important to consider the effect of technology change on pro-environmental behaviors. On the one hand, many organizations – governmental and non-governmental – spend considerable resources encouraging pro-environmental behaviors. On the other hand, technology change that is encouraged by similar or the same organizations has the potential to discourage these behaviors when that technology change reduces the environmental impact of consumption and therefore reduces the effectiveness of pro-environmental behaviors. If technology change makes sacrifices for the environment redundant then that negates the need to encourage pro-environmental behaviors. However, environmental policy challenges such achieving the massive emission cuts required to meet global climate change targets over the next few decades requires many actions – both technology change and sacrifice of consumption. Thus, it is important to consider how technology change interacts with incentives for pro-environmental behaviors to ensure resources expended on behavior change are allocated in an optimal way.

# References


Abbott, A., S. Nandeibam, and L. O'Shea (2013). Recycling: Social norms and warm-glow revisited. *Ecological Economics 90*, 10–18.

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics 95*(9-10), 1082–1095.

Allcott, H. and S. Mullainathan (2010). Behavior and energy policy. *Science 327*(5970), 1204–1205.

Araujo, F. A., E. Carbone, L. Conell-Price, M. W. Dunietz, A. Jaroszewicz, R. Landsman, D. Lam, L. Vesterlund, S. W. Wang, and A. J. Wilson (2016). The slider task: An example of restricted inference on incentive effects. *Journal of the Economic Science Association 2*(1), 1–12.

Ariely, D., A. Bracha, and S. Meier (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *The American Economic Review 99*(1), 544–555.

Ayres, I., S. Raseman, and A. Shih (2013). Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. *Journal of Law, Economics, and Organization 29*(5), 992–1022.

Bénabou, R. and J. Tirole (2006). Incentives and prosocial behavior. *The American Economic Review 96*(5), 1652–1678.

Bénabou, R. and J. Tirole (2011). Identity, morals, and taboos: Beliefs as assets. *The Quarterly Journal of Economics 126*(2), 805–855.

Benndorf, V., H. A. Rau, and C. Sölch (2018). Minimizing learning behavior in repeated real-effort tasks. *Discussion Papers, Center for European, Governance and Economic Development Research, No. 343*.

Berg, J., J. Dickhaut, and K. McCabe (1995). Trust, reciprocity, and social history. *Games and Economic Behavior 10*, 122–142.

Bertrand, M. and S. Mullainathan (2001). Do people mean what they say? Implications for subjective survey data. *American Economic Review 91*(2), 67–72.

Blanken, I., N. van de Ven, and M. Zeelenberg (2015). A meta-analytic review of moral licensing. *Personality and Social Psychology Bulletin 41*(4), 540–558.

Bowles, S. and S. Polania-Reyes (2012). Economic incentives and social preferences: Substitutes or complements? *Journal of Economic Literature 50*(2), 368–425.

Brekke, K. A., S. Kverndokk, and K. Nyborg (2003). An economic model of moral motivation. *Journal of Public Economics 87*(9-10), 1967–1983.

Cason, T. N. and L. Gangadharan (2002). Environmental labeling and incomplete consumer information in laboratory markets. *Journal of Environmental Economics and Management 43*(1), 113 – 134.

Chan, N. W. and K. Gillingham (2015). The microeconomic theory of the rebound effect and its welfare implications. *Journal of the Association of Environmental and Resource Economists 2*(1), 133–159.

Chang, J.-j., W.-N. Wang, and J.-Y. Shieh (2017). Environmental rebounds/backfires: Macroeconomic implications for the promotion of environmentally-friendly products. *Journal of Environmental Economics and Management*.

Costa, D. L. and M. E. Kahn (2013). Energy conservation nudges and environmentalist ideology: Evidence from a randomized residential electricity field experiment. *Journal of the European Economic Association 11*(3), 680–702.

Croson, R. and N. Treich (2014). Behavioral environmental economics: Promises and challenges. *Environmental and Resource Economics 58*(3), 335–351.

De Borger, B., I. Mulalic, and J. Rouwendal (2016). Measuring the rebound effect with micro data: A first difference approach. *Journal of Environmental Economics and Management 79*, 1–17.

DEFRA (2008). A framework for pro-environmental behaviours. Technical report, Department for Environment, Food and Rural Affairs, London.

Dinda, S. (2004). Environmental kuznets curve hypothesis: A survey. *Ecological Economics 49*(4), 431–455.

Duarte, R., V. Pinilla, and A. Serrano (2014). Looking backward to look forward: Water use and economic growth from a long-term perspective. *Applied Economics 46*(2), 212–224.

Dunlap, R. E., K. D. Van Liere, A. G. Mertig, and R. E. Jones (2000). Measuring endorsement of the New Ecological Paradigm: A revised NEP scale. *Journal of Social Issues 56*(3), 425–442.

Erkal, N., L. Gangadharan, and B. H. Koh (2017). Monetary and Non-Monetary Incentives in Real-Effort Tournaments. Available at SSRN: https://ssrn.com/abstract=3035906.

Erkal, N., L. Gangadharan, and N. Nikiforakis (2011). Relative earnings and giving in a real-effort experiment. *The American Economic Review 101*(7), 3330–3348.

Ferraro, P. J. and M. K. Price (2013). Using nonpecuniary strategies to influence behavior: Evidence from a large-scale field experiment. *Review of Economics and Statistics 95*(1), 64–73.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*(2), 171–178.

Friesen, L. and L. Gangadharan (2013). Environmental markets: What do we learn from the lab? *Journal of Economic Surveys 27*(3), 515–535.

Ghosh, N. K. and M. F. Blackhurst (2014). Energy savings and the rebound effect with multiple energy services and efficiency correlation. *Ecological Economics 105*, 55–66.

Gillingham, K., D. Rapson, and G. Wagner (2016). The rebound effect and energy efficiency policy. *Review of Environmental Economics and Policy 10*(1), 68–88.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association 1*(1), 114–125.

Halvorsen, B. (2008). Effects of norms and opportunity cost of time on household recycling. *Land Economics 84*(3), 501–516.

Hawcroft, L. J. and T. L. Milfont (2010). The use (and abuse) of the New Environmental Paradigm scale over the last 30 years: A meta-analysis. *Journal of Environmental Psychology 30*(2), 143–158.

Hunt, L. C. and D. L. Ryan (2014). Catching on the rebound: Why price elasticities are generally inappropriate measures of rebound effects. Technical report, Surrey Energy Economics Centre (SEEC), School of Economics, University of Surrey. Available at https://ideas.repec.org/p/sur/seedps/148.html.

Jevons, W. (1865). *The coal question: An inquiry concerning the progress of the nation, and the probable exhaustion of our coal-mines*. Macmillan.

Levitt, S. D. and J. A. List (2007). What do laboratory experiments measuring social preferences reveal about the real world? *The Journal of Economic Perspectives 21*(2), 153–174.

Monin, B. and D. T. Miller (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology 81*(1), 33.

Nyborg, K., R. B. Howarth, and K. A. Brekke (2006). Green consumers and public policy: On socially contingent moral motivation. *Resource and Energy Economics 28*(4), 351–366.

Schubert, C. (2017). Green nudges: Do they work? Are they ethical? *Ecological Economics 132*, 329–342.

Sorrell, S. and J. Dimitropoulos (2008). The rebound effect: Microeconomic definitions, limitations and extensions. *Ecological Economics 65*, 636–649.

Sorrell, S., J. Dimitropoulos, and M. Sommerville (2009). Empirical estimates of the direct rebound effect: A review. *Energy Policy 37*(4), 1356–1371.

Sturm, B. and J. Weimann (2006). Experiments in environmental economics and some close relatives. *Journal of Economic Surveys 20*(3), 419–457.

The Global Commission on the Economy and Climate (2014). Better growth better climate: The new climate economy report. Technical report, New Climate Economy.

Tiefenbeck, V., T. Staake, K. Roth, and O. Sachs (2013). For better or for worse? Empirical evidence of moral licensing in a behavioral energy conservation campaign. *Energy Policy 57*, 160–171.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). Cambridge, Mass: MIT Press.

# A Appendix

## A.1 Supplementary results

Table A.1: Tobits testing treatment effects, separated by environmental behaviors.

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Proportion pro-environmental effort | | | |
| | Envi beh. > median | | Envi beh. ≤ median | |
| | (1) | | (2) | |
| | Coefs | AMEs | Coefs | AMEs |
| Low damage | -0.0655* | -0.0305* | -0.1857* | -0.0598** |
| | (0.0347) | (0.0162) | (0.0953) | (0.0296) |
| Low effort | -0.0290 | -0.0136 | 0.0487 | 0.0164 |
| | (0.0522) | (0.0242) | (0.0513) | (0.0173) |
| Income effect | -0.0269 | -0.0126 | -0.0283 | -0.0094 |
| | (0.0341) | (0.0159) | (0.0758) | (0.0250) |
| Choice | -0.0247 | -0.0116 | 0.2803 | 0.0980 |
| | (0.0560) | (0.0261) | (0.2692) | (0.0972) |
| Chose low | 0.3454 | 0.1591 | -0.4239 | -0.1352 |
| | (0.3488) | (0.1516) | (0.4333) | (0.1299) |
| Choice*Chose low | -0.1507** | -0.0690** | -0.5137* | -0.1496** |
| | (0.0743) | (0.0334) | (0.2797) | (0.0682) |
| TG B | -0.5374** | -0.2189** | -0.1815 | -0.0580 |
| | (0.2678) | (0.0859) | (0.3709) | (0.1130) |
| TG C | -0.0345 | -0.0161 | 0.0113 | 0.0038 |
| | (0.2379) | (0.1106) | (0.2937) | (0.0982) |
| TG D | -0.4996 | -0.2111* | 0.6355 | 0.2302 |
| | (0.3464) | (0.1217) | (0.4704) | (0.1717) |
| TG E | -0.5171 | -0.2159 | 0.1562 | 0.0530 |
| | (0.4110) | (0.1413) | (0.4978) | (0.1708) |
| Constant | 0.3870** | | -0.0297 | |
| | (0.1521) | | (0.2147) | |
| $\hat{\sigma}$ | 0.4964*** | | 0.8617*** | |
| | (0.1022) | | (0.2053) | |
| N | | 101 | | 114 |
| Observations | | 303 | | 342 |
| P-value | | 0.0087 | | 0.0399 |
| Pseudo r-squared | | 0.0268 | | 0.0304 |
| Pseudo loglik | | -290.42 | | -318.91 |

Notes: TG abbreviates "Treatment group". Standard errors are clustered at the subject level and in parentheses. The delta-method is used to calculate standard errors for average marginal effects (AMEs). *p<0.1; **p<0.05; ***p<0.01.

Table A.2: OLS testing treatment effects as in Table 5.

| | *Dependent variable:* | |
|---|---|---|
| | Prop. pro-environmental effort | |
| | (1) | (2) |
| Low damage | -0.0501*** | -0.0499** |
| | (0.0169) | (0.0213) |
| Low effort | 0.0083 | -0.0069 |
| | (0.0159) | (0.0283) |
| Income effect | -0.0090 | -0.0278 |
| | (0.0129) | (0.0194) |
| Choice | 0.0564 | 0.0614 |
| | (0.0622) | (0.0886) |
| Chose low | -0.0439 | 0.0691 |
| | (0.1236) | (0.1262) |
| Choice*Chose low | -0.1311** | -0.1548* |
| | (0.0627) | (0.0851) |
| TG B | -0.1548* | -0.1379 |
| | (0.0843) | (0.0926) |
| TG C | -0.0096 | -0.0185 |
| | (0.0810) | (0.0855) |
| TG D | 0.0041 | -0.0944 |
| | (0.1261) | (0.1275) |
| TG E | -0.0459 | 0.1757 |
| | (0.1437) | (0.2511) |
| Low damage*TG B | | -0.0471 |
| | | (0.0404) |
| Low damage*TG C | | 0.0113 |
| | | (0.0385) |
| Low effort*TG C | | 0.0154 |
| | | (0.0420) |
| Low effort*TG D | | 0.0281 |
| | | (0.0378) |
| Inc. effect*TG E | | 0.0298 |
| | | (0.0262) |
| Choice*TG E | | 0.0045 |
| | | (0.1033) |
| Chose low*TG E | | -0.3681 |
| | | (0.2798) |
| Ch.*Ch. low*TG E | | 0.0422 |
| | | (0.1020) |
| Constant | 0.3704*** | 0.3754*** |
| | (0.0575) | (0.0596) |
| N | 215 | 215 |
| P-value | 0.0000 | 0.0000 |
| Ajd. r-squared | 0.0119 | 0.0081 |

Notes: TG abbreviates "Treatment group". Standard errors are clustered at the subject level and in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A.3: OLS of proportion pro-environmental effort in High damage treatment, treatment groups A, B, D, E as in Table 7.

| | *Dependent variable:* | | |
|---|---|---|---|
| | Proportion pro-environmental effort | | |
| | (1) | (2) | (3) |
| Norm belief | 0.2481*** | | 0.2299*** |
| | (0.0245) | | (0.0257) |
| Letters | | -0.0011 | 0.0004 |
| | | (0.0014) | (0.0012) |
| Age | | 0.0065 | 0.0016 |
| | | (0.0091) | (0.0074) |
| Female | | 0.1638*** | 0.0690 |
| | | (0.0576) | (0.0480) |
| Low income | | -0.1209 | -0.1111* |
| | | (0.0796) | (0.0647) |
| Australian | | 0.0555 | 0.0375 |
| | | (0.0742) | (0.0603) |
| Env. behav. | | 0.1419** | 0.0775 |
| | | (0.0653) | (0.0536) |
| NEP scale | | -0.0849 | -0.0533 |
| | | (0.0670) | (0.0546) |
| Env. org. | | 0.1464* | 0.1066 |
| | | (0.0883) | (0.0719) |
| Political party | | 0.3386** | 0.2061 |
| | | (0.1623) | (0.1327) |
| Vote Labor | | -0.1177 | -0.1090 |
| | | (0.1052) | (0.0855) |
| Vote Greens | | 0.1272 | 0.1407 |
| | | (0.1282) | (0.1042) |
| Vote other | | -0.0707 | -0.1508 |
| | | (0.1912) | (0.1556) |
| Vote unsure | | 0.0788 | 0.0065 |
| | | (0.0854) | (0.0698) |
| Constant | 0.0534 | 0.0409 | -0.0693 |
| | (0.0347) | (0.4293) | (0.3490) |
| N | 167 | 167 | 167 |
| Observations | 167 | 167 | 167 |
| P-value | 0.0000 | 0.0049 | 0.0000 |
| Adj. r-squared | 0.3803 | 0.1017 | 0.4069 |

Notes: The "vote" dummy variables are relative to Vote Liberal. *p<0.1; **p<0.05; ***p<0.01.

Table A.4: OLS testing treatment effects, separated by NEP and environmental behaviors, as in Tables 8 and A.1.

| | *Dependent variable:* Proportion pro-environmental effort | | | |
|---|---|---|---|---|
| | NEP > median | NEP ≤ median | E beh. > median | E beh. ≤ median |
| | (1) | (2) | (3) | (4) |
| Low damage | -0.0363 | -0.0625*** | -0.0311** | -0.0670** |
| | (0.0275) | (0.0208) | (0.0153) | (0.0288) |
| Low effort | 0.0331 | -0.0127 | -0.0069 | 0.0214 |
| | (0.0203) | (0.0245) | (0.0280) | (0.0169) |
| Income effect | -0.0185 | -0.0036 | -0.0118 | -0.0054 |
| | (0.0138) | (0.0211) | (0.0175) | (0.0190) |
| Choice | 0.1579 | -0.0217 | -0.0116 | 0.0955 |
| | (0.1141) | (0.0491) | (0.0427) | (0.0931) |
| Chose low | 0.1432 | -0.2516 | 0.1769 | -0.1947 |
| | (0.1021) | (0.1711) | (0.1268) | (0.1683) |
| Choice*Chose low | -0.2252** | -0.0604 | -0.0792* | -0.1549* |
| | (0.1130) | (0.0543) | (0.0466) | (0.0933) |
| TG B | -0.2933** | -0.0538 | -0.2498** | -0.0681 |
| | (0.1301) | (0.1170) | (0.1157) | (0.1216) |
| TG C | -0.0615 | -0.0096 | -0.0043 | -0.0074 |
| | (0.1280) | (0.1136) | (0.1214) | (0.1084) |
| TG D | -0.3277** | 0.3263* | -0.2465** | 0.1974 |
| | (0.1308) | (0.1786) | (0.1232) | (0.1806) |
| TG E | -0.2400 | 0.1216 | -0.2552 | 0.0823 |
| | (0.1682) | (0.1857) | (0.1692) | (0.1894) |
| Constant | 0.4502*** | 0.3290*** | 0.4284*** | 0.3148*** |
| | (0.1053) | (0.0675) | (0.0824) | (0.0802) |
| N | 102 | 113 | 101 | 114 |
| Observations | 306 | 339 | 303 | 342 |
| P-value | 0.0001 | 0.001 | 0.0002 | 0.0339 |
| Adj. r-squared | .0429 | .0313 | .0191 | .0188 |

Notes: TG abbreviates "Treatment group". Standard errors are clustered at the subject level and in parentheses. *p<0.1; **p<0.05; ***p<0.01.

## A.2 Experimental instructions and survey questions

What follows is the experimental instructions given to treatment groups B and E. Treatment groups A, C and D did not have an incentivized question for the survey. Treatment group C was given the low damage parameters in the example and practice round as they were given this treatment first. Otherwise the instructions are identical between treatment groups.

Figure A.1: Overview instructions.

# Instructions

**Overview**

Please note: all instructions will be read aloud by the experimenter.

Please turn off all cell phones for the duration of this experiment.

As part of today's experiment, you will be participating in an activity and a survey. The activity will consist of 3 rounds. There will also be a practice round at the start of the activity. You will receive detailed instructions before the practice round and before the actual rounds. We anticipate the total time for the experiment to be roughly one and a half hours.

As part of today's experiment, you will be earning money. You will be paid based on your effort and the decisions you make from one of the 3 rounds in the activity. The computer will randomly choose the round for which you will be paid. In addition to any earnings you might have in these tasks, you will be given $10 for participating in today's experiment. Any money that you earn in the experiment will be paid to you in cash at the end of the experiment. All payments will be rounded to the nearest 5c.

After you have completed the activity, we would like you to answer some questions about yourself in the survey. Please take your time and answer honestly and as accurately as possible. You will not be identified in any way. Your survey answers will only be used for this experiment and will only be used by the researcher(s) involved in this project.

We are about to begin the activity. Please listen carefully. It is important that you understand the instructions of the task properly. If you do not understand, you will not be able to participate effectively. The task will be explained with examples. Do not talk or discuss the activity with people around you. There will be opportunities to ask questions at the end of each set of instructions if you are unsure how to perform the activity. At any time during this experiment, please remain in your seat and follow the instructions of the experimenter. Also, do not look at any other person's responses at any time during this experiment.

Figure A.2: Practice round instructions part 1.

## PRACTICE ROUND Instructions

The activity consists of a word encoding task. The task involves correctly assigning numbers to up to 9 random letters that you are given. First, the task will be explained in detail. You will have the opportunity to ask questions. Then you will be given a practice round, followed by another opportunity to ask questions. Finally, you will be given the task for 3 rounds. Each round will last 8 minutes.

**Explanation:**

Near the bottom of each screen you will be given the full alphabet in a random order. Below each letter will be a number. Here is an example of what you will see:

| D | P | Q | E | G | O | U | I | T | S | R | V | F | N | B | C | H | W | A | M | Y | X | Z | K | L | J |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 59 | 28 | 25 | 91 | 87 | 11 | 20 | 26 | 49 | 34 | 61 | 35 | 99 | 85 | 82 | 63 | 96 | 83 | 38 | 76 | 88 | 67 | 29 | 46 | 17 | 72 |

In the centre of the screen, to the left, you will see 6 randomly selected letters. This is your "word". Below each letter is an empty box, as shown here:

| WORD : | E | B | J | K | O | M |
|---|---|---|---|---|---|---|
| CODE: | | | | | | |

In order to encode the "word" it is necessary to click on each box with your mouse and type the number associated with each letter. You may only use the numbers at the top of your keyboard, and *not* the numbers in the number pad at the side of your keyboard.

After you have completed a word it will be counted if you click OK with your mouse. The OK button is located at the bottom of the screen. If you click OK you will be given a new word to encode. The computer will not give you a new word until the word you have encoded is correct.

You will earn money for each word you successfully encode. The amount you have earned so far in a round will be displayed at the top of your screen.

For the practice round, it has been set so that you "earn" 60c for each word you encode. However, as this is a practice round, you will **not** actually be paid this money. Your earnings will be paid to you at the end of the experiment for one out of the 3 actual rounds. The round that will be paid out will be decided at random by the computer, so you need to treat all non-practice rounds as if they may be the one that will affect your final earnings. The one round that is paid out will be the same for everyone in this session.

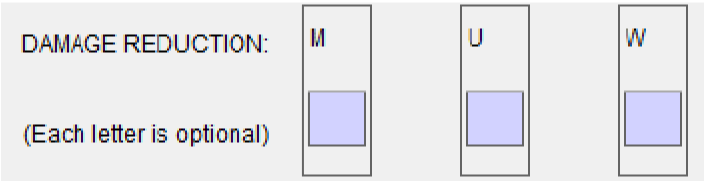
In addition to the earnings you make in a round, there will also be an amount donation to a charity. This amount starts at $336 for the full session per round, which is $14 per person per round. The one round that is chosen to be paid out for will also be the only round for which the donation to the charity is made.

Every word you encode will lower the charity payment. The damages per word to the charity has been set at 54c per word for the practice round. Again, this is for demonstration purposes only; the practice round will not be paid out. These damages mean that the first word you encode will lower the total session donation from $336 to $335.46. The second word you encode will lower the session donation

Figure A.3: Practice round instructions part 2.

by another 54c, down to $334.92, and so on. Remember, all other people in this room can lower the total session payment with their decisions. The total damages you have caused to the charity payment in each round will be displayed at the top of your screen, below your personal earnings.

You can reduce the damages to the charity. In the practice round you will be given 3 extra letters per word. These letters are **optional**. The extra letters are displayed to the right of the initial 6 letters that you must encode:

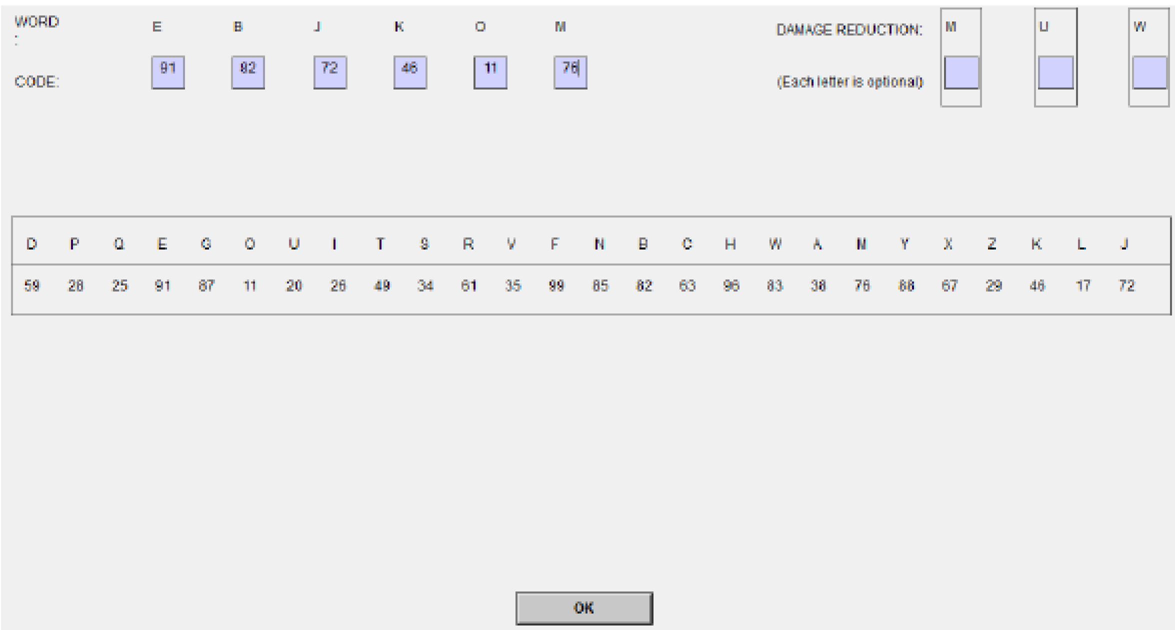



If you encode 1 of the optional extra letters, the damage to the charity for the current word you are encoding will reduce by 1 third, from 54c to 36c. You must encode this letter correctly for it to reduce the damages. If you encode 2 of the optional extra letters, the damage to the charity for the current word you are encoding will reduce by 2 thirds, from 54c to 18c. If you do not wish to cause any damage to the charity for the word you are currently encoding, you need to encode all three optional extra letters.

The money for the charity will be donated to a local environmental charity that plants indigenous trees in Victoria. Every $2 donated to the charity leads to one tree being planted. Therefore, every $2 of damages you cause to the charity's payment leads to 1 less tree being planted for that round. To prove the donation has been made, in the coming days you will receive an email with a receipt stating the total amount donated to the charity from everyone in this session. This information will include the average number of trees that will be planted per person from this session.

**Examples:**

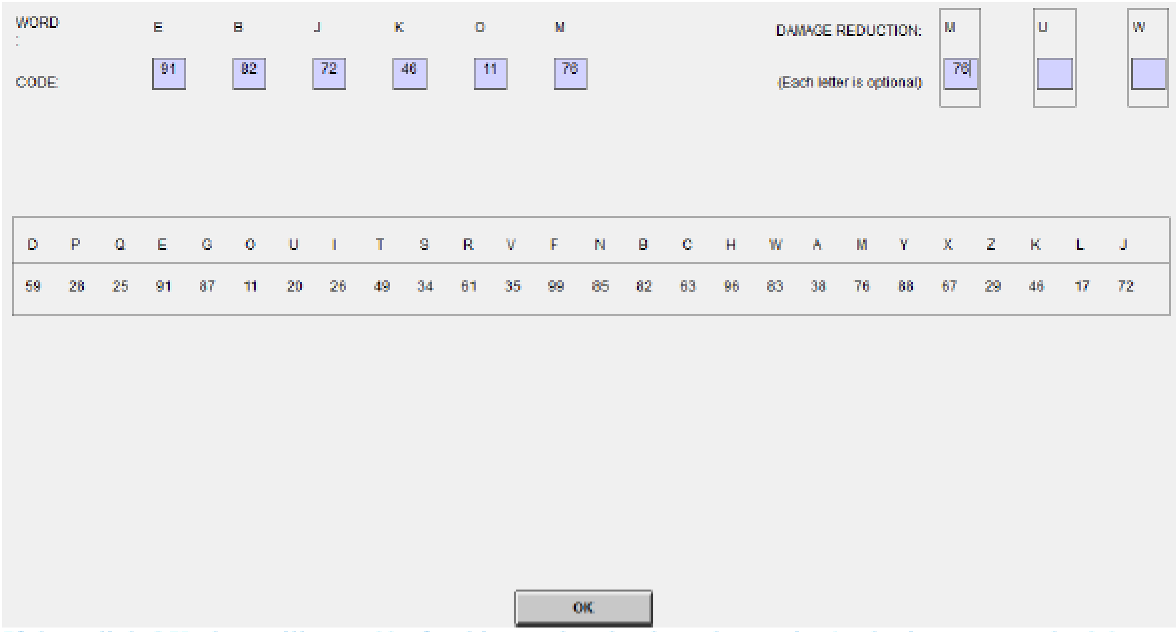
In the following image is an example of someone correctly encoding the word, and none of the optional extra letters:




They can now click OK. Once they have clicked OK, they will earn 60c and reduce the session's charity payment by 54c.

Figure A.4: Practice round instructions part 3.

In this next image, the person has added the correct code for one of the optional extra letters:




If they click OK, they will earn 60c for this word and reduce the session's charity payment by 36c.

**To recap:**

- You will be given a practice round. This round will not be paid out. Then you will be given 3 rounds of the task. One of these rounds will be chosen randomly by the computer to be paid out. Each round is 8 minutes (or 480 seconds) long.
- You can only use the numbers at the top of the keyboard, and you can only use your mouse to click on the boxes where you enter the numbers.
- You need to correctly encode the first 6 random letters to have a word counted. Each word earns 60c in the practice round, but given it is a practice round, these earnings will not be paid out.
- Each word you encode will cause 54c of damages to the charity in the practice round. For the word you are encoding, you can reduce the damages to 36c if you encode one of the optional extra letters, to 18c if you encode two of the optional extra letters, and to 0 if you encode three of the optional extra letters.
- Every $2 of damages you cause in the round that is paid out will lead to 1 less tree being planted.

*If you have any questions about this Activity please raise your hand now.*

**Quiz:**

You will now be given a quiz consisting of 4 quick questions to ensure you understand the activity. Feel free to refer to your printed instructions to find the correct answer.

You will earn 25c for each of the 4 questions you answer correctly. Please note that some questions have 2 parts.

The quiz will now begin.

Figure A.5: Practice round instructions part 4.

*After the quiz.*

**Practice round:**

We are about to begin the 8 minute practice round. The screen will appear exactly as it will in the actual activity. Note that the number of seconds remaining for the round is displayed at top right of the screen. Please use the practice round to familiarise yourself with the task, including how the earnings and damages are calculated.

The practice round will begin shortly.

Figure A.6: Activity instructions.

## ACTIVITY Instructions

*If you have any questions about the activity that have arisen from the practice round, please raise your hand now.*

You will now be given the task for 3 rounds. Each round will be 8 minutes long.

Remember, just one of the 3 rounds will be randomly chosen by the computer to be paid out. Therefore, you need to treat every round as though it will be paid out.

You will only be told how much will be donated to the charity at the end of session, based on which round is chosen to be paid out.

The payment amounts for each word, and the damages per word will change between the rounds. These amounts will be displayed on your screen before the start of each round, as well as be announced aloud by the experimenter.

*Are there any final questions?*

The first round will begin shortly.

Figure A.7: Survey instructions.

## SURVEY Instructions

You will now be given a survey. Please take your time to answer the questions truthfully.

Please note – for the second question in the survey, you will receive $1 if you answer it correctly. If you earn it, this additional $1 will be added to your final payment. The additional payment only applies to the second question. It will be clear which question the additional $1 payment applies to before you answer the question.

At the end of the survey, 1 out of the 3 rounds will be chosen to be paid out. You will be told which round will be paid out and therefore how much will be donated to the charity from this session. You will then receive your cash payment.

*Are there any questions?*

The survey will now begin.

Table A.5: Pro-environmental behaviours survey questions.

| | |
|---|---|
| Now you will be asked about some environmental behaviors. Please answer honestly - remember the answers to this survey are anonymous.<br><br>Thinking back over the past year, how often do you: | |
| 1. Take shorter showers to save water | Never / Rarely / Sometimes / Often / Always |
| 2. Turn off the tap when brushing your teeth | Never / Rarely / Sometimes / Often / Always |
| 3. Wash only full loads of clothes | Never / Rarely / Sometimes / Often / Always /NA |
| 4. Run the dishwasher only when full | Never / Rarely / Sometimes / Often / Always / NA |
| 5. Fix or report leaks when you notice them | Never / Rarely / Sometimes / Often / Always / NA |
| 6. Use the half flush button on the toilet when available | Never / Rarely / Sometimes / Often / Always |
| 7. Put rubbish in the bin | Never / Rarely / Sometimes / Often / Always |
| 8. Place cigarette butts in the bin | Never / Rarely / Sometimes / Often / Always / NA |
| 9. Recycle glass, hard plastics, paper and cans | Never / Rarely / Sometimes / Often / Always |
| 10. Use public transportation | Never / Rarely / Sometimes / Often / Always |
| 11. Ride a bicycle for transportation | Never / Rarely / Sometimes / Often / Always |
| 12. Choose to buy an organic option for a product if it is available | Never / Rarely / Sometimes / Often / Always |
| 13. Choose to buy a product on the basis that it has less packaging than a similar product | Never / Rarely / Sometimes / Often / Always |
| 14. Turn the lights off when leaving a room | Never / Rarely / Sometimes / Often / Always |
| 15. Use re-usable bags when shopping | Never / Rarely / Sometimes / Often / Always |

Table A.6: New Ecological Paradigm (NEP) survey questions (Dunlap et al., 2000).

| |
|---|
| Listed below are statements about the relationship between humans and the environment. For each one, please indicate whether you STRONGLY AGREE, MILDLY AGREE, are UNSURE, MILDLY DISAGREE or STRONGLY DISAGREE with it. |
| 1. We are approaching the limit of the number of people the earth can support.<br>2. Humans have the right to modify the natural environment to suit their needs.<br>3. When humans interfere with nature it often produces disastrous consequences.<br>4. Human ingenuity will insure that we do NOT make the earth unlivable.<br>5. Humans are severely abusing the environment.<br>6. The earth has plenty of natural resources if we just learn how to develop them.<br>7. Plants and animals have as much right as humans to exist.<br>8. The balance of nature is strong enough to cope with the impacts of modern industrial nations.<br>9. Despite our special abilities humans are still subject to the laws of nature.<br>10. The so-called ecological crisis facing humankind has been greatly exaggerated.<br>11. The earth is like a spaceship with very limited room and resources.<br>12. Humans were meant to rule over the rest of nature.<br>13. The balance of nature is very delicate and easily upset.<br>14. Humans will eventually learn enough about how nature works to be able to control it.<br>15. If things continue on their present course, we will soon experience a major ecological catastrophe. |